Proceedings of the 5th international conference on disability, virtual reality and associated technologies (ICDVRAT 2004)

Book

Published Version

Conference Proceedings

Sharkey, P., McCrindle, R. and Brown, D., eds. (2004) Proceedings of the 5th international conference on disability, virtual reality and associated technologies (ICDVRAT 2004). ICDVRAT. The University of Reading, Reading, UK, pp344. ISBN 0704911442 Available at http://centaur.reading.ac.uk/15090/

Publisher: The University of Reading

The 5th International Conference on

Disability, Virtual Reality and Associated Technologies

Proceedings

Edited by:

Paul Sharkey (Programme Chair)
Rachel McCrindle (Conference Co-Chair)
David Brown (Conference Co-Chair)

20 to 22 of September, 2004

Oxford, UK

ICDVRAT 2004

Proceedings:

The papers appearing in this book comprise the proceedings of the 5th International Conference on Disability, Virtual Reality and Associated Technologies, held on the 20th, 21st and 22nd of September, 2004 in New College, Oxford, UK. The papers presented reflect the authors' opinions and are published as presented and without change (formatting and minor editing excepted). Their inclusion in this publication does not necessarily constitute endorsement by the editors, by ICDVRAT, or by the University of Reading.

Please use the following format to cite material from these Proceedings:

A.B. Author(s), "Title of paper", *Proc. 5th Intl Conf. on Disability, Virtual Reality and Assoc. Technologies,* in Sharkey, McCrindle & Brown (Eds.), page numbers, Oxford, UK, 20–22 Sept. 2004.

Proceedings reference number: ISBN 07 049 11 44 2

Published by

The University of Reading

For information, contact:
ICDVRAT, Department of Cybernetics, University of Reading,
Whiteknights, Reading, RG6 6AY, UK

Phone: +44 (0) 118 378 6704 Fax: +44 (0) 118 378 8220

Email: *p.m.sharkey@reading.ac.uk* Web: *www.cyber.rdg.ac.uk/icdvrat/*

Printed in the UK.

Contents

Abstracts & Information on Oxford

Session I Stroke Rehabilitation

Session Chair: Rachel McCrindle

Session II Training and Virtual Transport

Session Chair: Hideyuki Sawada

Session III Virtual Environments and Games for Special Needs

Session Chair: William Harwin

Session VI Audio Virtual Environments

Session Chair: Tomohiro Kuroda

Session VII Brain Injury and Rehabilitation

Session Chair: Craig Carignan

Session VIII Speech and Communication

Session Chair: Jaime Sánchez

Session IX Interfacing to and Navigation within Virtual Environments

Session Chair: Tony Lewis-Brooks

Session X Virtual Environments for Assessment

Session Chair: David Brown

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

Conference Organisation

Conference Co-Chairs

Dr Rachel McCrindle, University of Reading, UK
Dr David Brown, Nottingham Trent University, UK

Programme Chair

Professor Paul Sharkey, University of Reading, UK

International Programme Committee

Mr Jurgen Broeren, Sahlgrenska Academy, Göteborg University, Sweden
Dr Jane Broida, Metropolitan State College of Denver, USA
Dr Barbara Brooks, University of East London, UK
Dr Sue Cobb, University of Nottingham, UK
Mr Martyn Cooper, Open University, UK
Professor Kerstin Dautenhahn, University of Hertfordshire, UK
Dr Roy Davies, Lund University, Sweden
Dr Alistair Edwards, University of York, UK
Mr Michael Grant, University of Strathclyde, UK
Dr William Harwin, University of Reading, UK
Dr Colin Harrison, University of Strathclyde, UK
Professor Gunnar Jansson, Uppsala Universitet, Sweden
Professor Gerd Johansson, Lund University, Sweden
Prof Noomi Katz, The Hebrew University of Jerusalem, Israel
Dr Tomohiro Kuroda, Kyoto University Hospital, Japan
Professor Peter Kyberd, University of New Brunswick, Canada
Dr Sarah Nichols, University of Nottingham, UK
Professor Helen Petrie, City University, London, UK
Dr Luigi Pugnetti, Fondazione Don Gnocchi, Milan, Italy
Professor Albert Rizzo, University of Southern California, Los Angeles, USA
Dr David Roberts, University of Salford, UK
Professor David Rose, University of East London, UK
Dr Cecília Sik Lányi, University of Veszprém, Hungary
Dr Jaime Sanchez, Universidad de Chile, Chile
Dr Penny Standen, University of Nottingham, UK
Dr Danaë Stanton, University of Bath, UK
Dr Brenda Wiederhold, California School of Professional Psychology, USA
Dr Mark Wiederhold, CyberPsychology and Behavior Journal
Dr Paul Wilson, University of Hull, UK
Professor John Wilson, University of Nottingham, UK

Local Organising Committee

Dr Rachel McCrindle, University of Reading, UK (Chair)
Professor Paul Sharkey, University of Reading, UK
Mr Richard Sherwood, ICDVRAT *Web Manager*, University of Reading, UK
Mr Aidan Caffrey, University of Reading, UK
Mr Mark Foyle, University of Reading, UK

Introduction

The purpose of the 5th International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT 2004) is to provide a forum for international experts, researchers and user groups to present and review how advances in the general area of Virtual Reality can be used to assist people with Disability.

After a peer review process, the International Programme Committee selected 34 papers for presentation at the conference, collected into 10 plenary sessions: Stroke Rehabilitation; Training and Virtual Transport; Virtual Environments and Games for Special Needs; Enhancing Mobility and Accessibility; Haptics; Virtual Environments for Assessment; Brain Injury and Rehabilitation; Speech and Communication; Interfacing to and Navigation within Virtual Environments and Audio Virtual Environments. There is an additional session specifically for informal demonstrations, poster presentations and exhibits from a small number of companies.

The conference will be held over three days at New College, in Oxford, UK.

ICDVRAT 2004 marks the 10th year of the conference series, following on from the success of conferences held in Maidenhead, UK (1996), Skövde, Sweden (1998), Alghero, Sardinia, Italy (2000) and Veszprém, Hungary (2002). Abstracts from this conference and full papers from the previous conferences are available online from the conference web site *www.icdvrat.reading.ac.uk*. We are also pleased to be able to provide the complete ICDVRAT archive on CD-ROM with this Proceedings.

Acknowledgements

The Conference Chairs would like to thank the Programme Committee, for their input to the conference format and focus, and for their commitment to the review process, the authors of all the papers submitted to the conference, the Organisation Committee, and the students who help out over the period of the conference. On behalf of ICDVRAT 2004, we welcome all delegates to the Conference and sincerely hope that delegates find the conference to be of great interest.

The Conference Chairs for ICDVRAT welcome any feedback on this year's conference. We would also welcome suggestions for the venue/host country for ICDVRAT 2006.

Paul Sharkey, Rachel McCrindle and David Brown

 ISBN 07 049 11 44 2

Conference Sponsors

ICDVRAT 2004 is sponsored by the Department of Cybernetics, School of Systems Engineering, of the University of Reading, UK.

Additional help in publicising the conference has been gratefully received from vrpsych-l@usc.edu and uk-vrsig@mailbase.ac.uk, amongst many others.

Artwork

Artwork and Conference Layout by sckelp, adapted from original ideas of Eric Phipps and David Angus (ICDVRAT Logo) and from original clipart from the CorelDraw Clipart Library (ICDVRAT 2004 Graphic). Eric Phipps and David Angus (both of Project DISCOVER) and sckelp may be contacted through ICDVRAT.

Abstracts

In alphabetical order, based on first author.

Survey of modelling approaches for medical simulators, **A Al-khalifah** and **D Roberts**, University of Reading/University of Salford, Manchester, UK

Medical simulation, in particular that used for training and planning, has become an established application of Virtual Reality technology. This application became an active area of simulation development in many academic and commercial institutions around the globe. A reasonable number of successful commercial medical simulators have been launched, while others remain hostage in research laboratories undergoing further developments and improvements. This paper provides a dichotomy of modelling techniques in the context of deformation and cutting, giving examples of how these are applied in medical simulation, comparing their strengths and weaknesses, outlining limitations and pinpoint expectations for the future. We focus on mapping the aim of the simulator to the adoption of particular modelling approaches. A case study pays special attention to the simulation of human organs where we uncover advances and limitations in the application of these modelling approaches.

Using switch controlled software with people with profound disabilities, **N Anderton**, **P J Standen** and **K Avory**, University of Nottingham, UK

Micro switches used to control sound and light displays increase the level of activity of people with profound intellectual disabilities and provide a means by which they can exert some control over their environments. This study set out to i) explore whether people with profound disabilities could learn to use a simple game controlled by a single micro switch and displayed on a normal computer monitor; ii) document what activities on the part of a tutor best facilitated the performance of the learner. Four men and three women aged between 24 and 46 years with profound disabilities completed eight twice weekly sessions when they were given the opportunity to play a computer game that could be operated by a large jelly bean switch. A tutor sat next to them throughout the sessions and each session was recorded on videotape. Tapes were analysed for the help given by the tutor, use of the switch and duration of attention. Although the game was too difficult for them, all participants increased the percentage of time during the session in which they looked at the monitor and for all of them there were at least two sessions when their switch pressing became a consequence of the tutor's activity.

Towards eye based virtual environment interaction for users with high-level motor disabilities, **R Bates** and **H O Istance**, De Montfort University, Leicester, UK

An experiment is reported that extends earlier work on the enhancement of eye pointing in 2D environments, through the addition of a zoom facility, to its use in virtual 3D environments using a similar enhancement. A comparison between hand pointing and eye pointing without any enhancement shows a performance advantage for hand based pointing. However, the addition of a 'fly' or 'zoom' enhancement increases both eye and hand based performance, and reduces greatly the difference between these devices. Initial attempts at 'intelligent' fly mechanisms and further enhancements are evaluated.

Design, development and manufacture of novel assistive and adaptive technology devices, **S J Battersby**, **D J Brown**, **P J Standen**, **N Anderton** and **M Harrison**, Nottingham Trent University/University of Nottingham,/The Portland Partnership, Nottingham, ENGLAND

The aim of this research is to design, develop, evaluate and manufacture an assistive/adaptive computer peripheral to facilitate interaction and navigation within Virtual Learning Environments and related learning content for people with physical learning and disabilities. The function of the device will be software specific; however the most common primary functions are those of selection, navigation and input.

Developing a multimodal web application, **A Caffrey** and **R J McCrindle**, University of Reading, UK

This paper describes the creation of a multi-modal website that incorporates both haptics and speech recognition. The purpose of the work is to provide a new and improved method of internet navigation for visually impaired users. The rationale for implementing haptic devices and speech recognition software within websites is described, together with the benefits that can accrue from using them in combination. A test site has been developed which demonstrates, to visually impaired users, several different types of web application that could make use of these technologies. It has also been demonstrated that websites incorporating haptics and speech recognition can still adhere to standard usability guidelines such as Bobby. Several tests were devised and undertaken to gauge the effectiveness of the completed web site. The data obtained has been analysed and provides strong evidence that haptics and speech recognition can improve internet navigation for visually impaired users.

Time-scale modification as a speech therapy tool for children with verbal apraxia, **E Coyle**, **O Donnellan**, **E Jung**, **M Meinardi**, **D Campbell**, **C MacDonaill** and **P K Leung**, Dublin Institute of Technology, IRELAND

A common suggested treatment for verbal apraxia is repetition, and the use of slow speech. The required slow speech may be attained by time-scaling ordinary-speed speech. However, when used for this purpose, the quality of the expanded speech must be of a very high quality to be of pedagogical benefit. This paper describes a new method of time-scaling based on the knowledge of speech characteristics, the relative durations of speech segments, and the variation of these durations with speaking rate. The new method achieves a high quality output making it suitable for use as a computer-assisted speech therapy tool.

Virtual reality in the rehabilitation of the upper limb after stroke: the user's perspective, **J H Crosbie**, **S M McDonough**, **S Lennon**, **L Pokluda** and **M D J McNeill**, University of Ulster, NORTHERN IRELAND

Virtual reality provides a three-dimensional computer representation of a real world or imaginary space through which a person can navigate and interact with objects to carry out specific tasks. One novel application of VR technology is in rehabilitation following stroke, particularly of the upper limb. Our research group has built a system for use in this field, which gives the user the ability to interact with objects by touching, grasping and moving their upper limb. A range of user perspectives has been tested with healthy individuals and with people following stroke.

Mixed reality environments in stroke rehabilitation: development as rehabilitation tools, **J A Edmans**, **J Gladman**, **M Walker**, **A Sunderland**, **A Porter** and **D Stanton Fraser**, University of Nottingham/University of Bath, UK

A virtual or mixed reality environment for neurological rehabilitation should simulate the rehabilitation of the task, and not simply simulate the task. This involves identifying the errors in task performance that neurologically damaged patients make during task performance and replicating the guidance given during skilled rehabilitation. Involvement of a skilled therapist in the design and development team is essential. Neurological rehabilitation is complex and replicating it requires compromises between the desire to replicate this complex process and the amount of development time. Virtual or mixed reality systems that can simulate the rehabilitation process are suitable for clinical effectiveness studies.

Computer games for children with visual impairments, **Y Eriksson** and **D Gärdenfors**, Göteborg University/ Royal Institute of Technology, Stockholm, SWEDEN

The Swedish Library of Talking Books and Braille (TPB) has published web-based computer games for children with different kinds of visual impairments. As the target groups have very different needs when it comes to the use of graphics and sound, TPB have developed two kinds of games. Image-based games aim to encourage children with partial sight to practise recognising visual objects, while sound-based games also intend to be accessible without relying on vision. Based on the results of two pilot studies, this paper discusses central design issues of the graphical and sound-based interfaces for this type of applications.

Promoting research and clinical use of haptic feedback in virtual environments, **U Feintuch**, **D Rand**, **R Kizony** and **P L Weiss**, Hadassah-Hebrew University Medical Center/University of Haifa/Beit-Rivka Geriatric Center, Petach-Tikva, ISRAEL

Converging evidence demonstrates the important role played by haptic feedback in virtual reality-based rehabilitation. Unfortunately many of the available haptic systems for research and intervention are rather costly, rendering them inaccessible for use in the typical clinical facility. We present a versatile and easy-to-use software package, based on an off-the-shelf force feedback joystick. We propose that this tool may be used for a wide array of research and clinical applications. Two studies, involving different populations and different applications of the system, are presented in order to demonstrate its usability for haptic research. The first study investigates the role of haptic information in maze solving by intact individuals, while the second study tests the usability of haptic maps as a mobility aid for children who are blind.

Interaction via motion observation, **M A Foyle** and **R J McCrindle**, University of Reading, UK

The main method of interacting with computers and consumer electronics has changed very little in the past 20 years. This paper describes the development of an exciting and novel Human Computer Interface (HCI) that has been developed to allow people to interact with computers in a visual manner. The system uses a standard computer web camera to watch the user and respond to movements made by the user's hand. As a result, the user is able to operate the computer, play games or even move a pointer by waving their hand in front of the camera. Due to the visual tracking aspect of the system, it is potentially suitable for disabled people whose condition may restrict their ability to use a standard computer mouse. Trials of the system have produced encouraging results, showing the system to have great potential as an input medium. The paper also discusses a set of applications developed for use with the system, including a game, and the implications such a system may have if introduced into everyday life.

New accessibility model for Microsoft windows, **R Haverty**, Microsoft Corporation, Washington, USA

Microsoft® Windows® User Interface (UI) Automation is the new accessibility framework for Microsoft Windows and is intended to address the needs of assistive technology products and automated testing frameworks by providing programmatic access to information about the user interface. UI Automation will be fully supported in the Windows platform on "Longhorn" and will be the means of enabling automated testing and accessibility for all new forms of Windows user interface, including existing legacy controls.

Using physiological measures for emotional assessment: a computer-aided tool for cognitive and behavioural therapy, **B Herbelin**, **P Benzaki**, **F Riquier**, **O Renault** and **D Thalmann**, Swiss Federal Institute of Technology, Lausanne/Adult Psychiatry University Department, Prilly, SWITZERLAND

In the context of Cognitive and Behavioural Therapies, the use of immersion technologies to replace classical exposure could improve the therapeutic process. As it is necessary to validate the efficiency of such a technique, both therapists and VR specialists need tools to monitor the impact of Virtual Reality Exposure on the patients. According to previous observations and experiments, it appears that an automatic evaluation of the Arousal and Valence components of affective reactions can provide significant information. The present study investigates a possible solution of Arousal and Valence computation from physiological measurements. Results show that the dimensional reduction is not statistically meaningful, but the correlations found encourage the investigation of this approach as a complement to cognitive and behavioural study of the patient.

Interactive virtual environment training for safe street crossing of right hemisphere stroke patients with unilateral spatial neglect, **N Katz**, **H Ring**, **Y Naveh**, **R Kizony**, **U Feintuch** and **P L Weiss**, Hadassah-Hebrew University/Tel Aviv University/University of Haifa, ISRAEL

The goal of this study was to determine whether non immersive interactive virtual environments are an effective medium for training individuals who suffer from Unilateral Spatial Neglect (USN) as a result of a right hemisphere stroke. Participants included 19 patients with stroke in two groups, an experimental group we were given VR-based street crossing training and a control group who were given computer based visual scanning tasks, both for a total of twelve sessions over four weeks. The results achieved by the VR street crossing intervention equalled those achieved by conventional visual scanning tasks. For some measures, the VR intervention even surpassed the scanning tasks in effectiveness. Despite several limitations in this study the present results support further development of the program.

Employing a virtual environment in postural research and rehabilitation to reveal the impact of visual information, **E A Keshner**, **R V Kenyon**, **Y Dhaher** and **J W Streepey**, Rehabilitation Institute of Chicago/Northwestern University, Chicago/University of Illinois at Chicago, USA

We have united an immersive virtual environment with support surface motion to record biomechanical and physiological responses to combined visual, vestibular, and proprioceptive inputs. We have examined age-related differences during peripheral visual field motion and with a focal image projected on to the moving virtual scene. Our data suggest that the postural response is modulated by all existing sensory signals in a non-additive fashion. An individual's perception of the sensory structure appears to be a significant component of the postural response in these protocols. We will discuss the implications of these results to clinical interventions for balance disorders.

Virtual reality based intervention in rehabilitation: relationship between motor and cognitive abilities and performance within virtual environments for patients with stroke, **R Kizony**, **N Katz** and **P L Weiss**, Hadassah-Hebrew University, Jerusalem/University of Haifa, Haifa, ISRAEL

The objective of this study was to provide experimental data to support a proposed model of VR-based intervention. More specifically our goal was to examine the relationships between cognitive and motor ability and performance within virtual environments. Thirteen participants who have had a stroke participated in the study. They each experienced three virtual environments (Birds & Balls, Soccer and Snowboard) delivered by the GX- video capture system. After each environment they complete a scenario specific questionnaire and Borg's scale for perceived exertion. Their cognitive, motor and sensory abilities were measured as well. The participants' responses to the VR environments showed that they enjoyed the experience and felt high levels of presence. The results also revealed some moderate relationships between several cognitive abilities and VR performance. In contrast, the motor abilities and VR performance were inversely correlated. In addition, there was a relationship between presence and performance within the Soccer environment. Although these results support some components of the proposed model it appears that the dynamic nature of the virtual experiences would be more suited to comparisons with different measures of motor ability than those used in the current study.

Design and user evaluation of a spatial audio system for blind users, **S H Kurniawan**, **A Sporka**, **V Nemec** and **P Slavik**, UMIST, Manchester, UK/Czech Technical University of Prague, CZECH REPUBLIC

The paper reports on the design and evaluation of a spatial audio system that models the acoustic response of a closed environment with varying sizes and textures. To test the fit of the algorithms used, the system was evaluated by nine blind computer users in a controlled experiment using seven distinct sounds in three environments. The statistical analysis reveals that there was insignificant difference in user perception of room sizes between sounds in real and simulated scenes. This system can contribute to the area of VR systems used for training blind people to navigate in real environments.

Simulating the Effects of Fatigue and Pathologies in Gait, **T Komura**, **A Nagano**, **H Leung** and **Y Shinagawa**, City University of Hong Kong, HONG KONG/RIKEN, Saitama, JAPAN/University of Illinois at Urbana-Champaign, USA

In this study, we propose a new method to simulate the effect of fatigue and pathologies in human gait motion. The method is based on Angular Momentum inducing inverted Pendulum Mode (AMPM), which is the enhanced version of 3D linear inverted pendulum mode that is used in robotics to generate biped locomotion. By importing the gait motion captured using a motion-capture device, the value of AMPM-parameters that define the trajectory of the center of mass and the angular momentum are calculated. By minimizing an objective function that takes into account the fatigue and disabilities of muscles, the original motion is converted to a new motion. Since the number of parameters to describe the motion is small in our method, the optimization process converges much more quickly than in previous methods.

Consumer price data-glove for sign language recognition, **T Kuroda**, **Y Tabata**, **A Goto**, **H Ikuta** and **M Murakami**, Kyoto University Hospital/Kyoto College of Medical Technology/AMITEQ Corp., Tokyo/Teiken Limited, Osaka, JAPAN

A data-glove available for full degrees of freedom of a human hand is a key device to handle sign language on information systems. This paper presents an innovative intelligent data-glove named StrinGlove. StrinGlove obtains full degrees of freedom of human hand using 24 Inductcoders and 9 contact sensors, and encodes hand postures into posture codes on its own DSP. Additionally, the simple structure of the glove decreases the price. Several sign experts tried the prototype and the results show that the prototype has sufficient recognition rate as a sensor unit and sufficient comfortableness as a glove to wear.

Blind Persons' Acquisition of Spatial Cognitive Mapping and Orientation Skills Supported by Virtual Environment, **O Lahav** and **D Mioduser**, Tel Aviv University, ISRAEL

Mental mapping of spaces, and of the possible paths for navigating these spaces, is essential for the development of efficient orientation and mobility skills. Most of the information required for this mental mapping is gathered through the visual channel. Blind people lack this crucial information and in consequence face great difficulties (a) in generating efficient mental maps of spaces, and therefore (b) in navigating efficiently within these spaces. The work reported in this paper follows the assumption that the supply of appropriate spatial information through compensatory sensorial channels, as an alternative to the (impaired) visual channel, may contribute to the mental mapping of spaces and consequently, to blind people's spatial performance. The main tool in the study was a virtual environment enabling blind people to learn about real life spaces, which they are required to navigate.

Evaluation of a computer-assisted 2D interactive virtual reality system in training street survival skills of people with stroke, **Y S Lam**, **S F Tam**, **D W K Man** and **P L Weiss**, Kong Polytechnic University, HONG KONG/University of Haifa, ISRAEL

The rationale, procedures and results of pilot study of the VR system in training street survival skills of people with stroke were presented and discussed. The following main study was also refined from the outcomes to make it more feasible and potentially beneficial to the patients.

HUMANICS 1 – a feasibility study to create a home internet based telehealth product to supplement acquired brain injury therapy, **A Lewis-Brooks**, Aalborg University Esbjerg, DENMARK

The goal to produce a unique, cost effective, and user-friendly computer based telehealth system product which had longevity and the ability to be integrated modularly into a future internet-based health care communication provision was conceptualised as an aid to home-based self-training. This through motivated creativity with the manipulation of multimedia. The system was to be a supplementary tool for therapists. The targeted group was initially to be those with acquired brain injury. This paper details phase 1 of the product feasibility testing.

Interactive Painting – an evolving study to facilitate reduced exclusion from classical music concerts for the deaf community, **A Lewis-Brooks**, Aalborg University Esbjerg, DENMARK

Exclusion from the joy of experiencing music, especially in concert venues, is especially applicable to those with an auditory impairment. There have been limited investigation into how to reduce the exclusion for this community in attending classical orchestra music concerts. Through utilizing computer technology and human machine interfaces (sensors and cameras) to stimulate complementary senses through interpretation it is possible to reduce this exclusion. Case studies are presented where the visual and tactile interpretation of the music is able to give new meaning and understanding for such people.

Creating aesthetically resonant environments for the handicapped, elderly and rehabilitation: Sweden, **A Lewis-Brooks** and **S Hasselblad**, Aalborg University Esbjerg, DENMARK/Emaljskolan, Landskrona, SWEDEN

This contribution expounds on our prior research, where interactive audiovisual content was shown to support Aesthetic Resonant Environments with (1) brain damaged children -extended here with the addition of (2) learning disabled, Parkinson's disease, and the aged. This paper appraises the experiments involved in preparing, developing and authenticating 'aesthetic resonance' within the Swedish partners' research (1 & 2). It reports on the inductive strategies leading to the development of the open architectural algorithms for motion detection, creative interaction and analysis, including the proactive libraries of interactive therapeutic exercise batteries based on multimedia manipulation in real-time.

Robot aided therapy: challenges ahead for upper limb stroke rehabilitation, **R C V Loureiro**, **C F Collin** and **W S Harwin**, University of Reading, UK

People who have been discharged from hospital following a stroke still have a potential to continue their recovery by doing therapy at home. Unfortunately it is difficult to exercise a stroke affected arm correctly and many people simply resort to using their good arm for most activities. This strategy makes many tasks difficult and any tasks requiring two hands become nearly impossible. The use of haptic interface technologies will allow the reach and grasp movements to be retrained by either assisting movement, or directing movement towards a specified target. This paper demonstrates how initial work on machine mediated therapies can be made available to a person recovering at home.

Participants Responses to a Stroke Training Simulator, **M Maxhall**, **A Backman**, **K Holmlund**, **L Hedman**, **B Sondell** and **G Bucht**, Umeå University, SWEDEN

The primary goal of this research was to study a virtual environments (VE) possibility to influence empathy on caregiver personal. In the present explorative study, 9 subjects from Norrlands University Hospital (NUS) completed a test consistent of three everyday tasks, reading a newspaper, filling a glass of water and putting toothpaste on a toothbrush. The procedure was done twice first from a non-stroke perspective and secondly from a perspective of a patient with stroke handicaps. The VE looked like a normal apartment and could be experienced with or without different perceptual disorders of stroke. Data from interviews and observations was analyzed via methods inspired by Grounded Theory. Results from observations and interviews indicate that the simulator in spite of problems of usability were effective in influencing caregivers empathy.

Development of a virtual reality system to study tendency of falling among older people, **L Nyberg**, **L Lundin-Olsson**, **B Sondell**, **A Backman**, **K Holmlund**, **S Eriksson**, **M Stenvall**, **E Rosendahl**, **M Maxhall** and **G Bucht**, Umeå University, SWEDEN

Injuries related to falls are a major threat to older persons health. A fall may not only result in an injury, but also in a decreased sense of autonomy in the persons daily life. In order to be able to prevent such falls there is a need to further understand the complex mechanisms involved in balance and walking. Here we present an immersive virtual reality system in which a person can move around, while being subjected to various events, which may influence balance and walking.

Cooperative control of virtual objects using haptic teleoperation over the internet, **A P Olsson**, **C R Carignan** and **J Tang**, Royal Institute of Technology, Stockholm, SWEDEN/Georgetown University, Washington, DC, USA

The feasibility of performing remote assessment and therapy of patients over the internet using robotic devices is explored. Using a force feedback device, the therapist can assess the range of motion, flexibility, strength, and spasticity of the patient's arm grasping a similar robotic device at a remote location. In addition, cooperative rehabilitation strategies can be developed whereby both the patient and therapist cooperatively perform tasks in a virtual environment. To counter the destabilizing effects of time delay in the force feedback loop, a passive wave variable architecture is used to encode velocity and force information. The control scheme is validated experimentally over the internet using a pair of InMotion2 robots located 500 miles apart.

Virtual environments as an aid to the design and evaluation of home and work settings for people with physical disabilities, **O Palmon**, **R Oxman**, **M Shahar** and **P L Weiss**, University of Haifa/Technion, ITT, ISRAEL

One of the major challenges facing the professionals involved in the home modification process is to succeed in adapting the environments in a way that enables an optimal fit between the individual and the setting in which he or she operates. The challenge originates primarily from the fundamental characteristic of design - one can see and test the final result of home modifications only after they have been completed. The goal of this study was to address this problem by developing and evaluating an interactive living environments model, HabiTest, that will facilitate the planning, design and assessment of optimal home and work settings for people with physical disabilities. This paper describes the Habitest tool, an interactive model that has been implemented via an immersive virtual reality system which displays three-dimensional renderings of specific environments, and which responds to user-driven manipulations such as navigation within the environment and alteration of its design. Initial results of a usability evaluation of this interactive environment by users are described.

Synthesis of virtual reality animation from sign language notation using MPEG-4 body animation parameters, **M Papadogiorgaki**, **N Grammalidis**, **N Sarris** and **M G Strintzis**, Informatics and Telematics Institute, Thermi-Thessaloniki/Olympic Games Organizing Committee, Athens 2004, GREECE

This paper presents a novel approach for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Face and Body Animation. Sign Language notation, in the well-known Sign Writing system, is provided as input and is initially converted to SWML (Sign Writing Markup Language), an XML-based format that has recently been developed for the storage, indexing and processing of Sign Writing notation. Each basic sign, namely sign box, is then converted to a sequence of Body Animation Parameters (BAPs) of the MPEG-4 standard, corresponding to the represented gesture. In addition, if a sign contains facial expressions, these are converted to a sequence of MPEG-4 Facial Animation Parameters (FAPs), while exact synchronization between facial and body movements is guaranteed. These sequences, which can also be coded and/or reproduced by MPEG-4 BAP and FAP players, are then used to animate H-anim compliant VRML avatars, reproducing the exact gestures represented in the sign language notation. Envisaged applications include interactive information systems for the persons with hearing disabilities (Web, E-mail, info–kiosks) and automatic translation of written texts to sign language (e.g. for TV newscasts).

Mixed reality environments in stroke rehabilitation: interfaces across the real/virtual divide, **T Pridmore**, **D Hilton**, **J Green**, **R Eastgate** and **S Cobb**, University of Nottingham, UK

Previous studies have examined the use of virtual environments (VEs) for stroke and similar rehabilitation. To be of real benefit it is essential that skills (re-)learned within a VE transfer to corresponding real-world situations. Many tasks have been developed in VEs, but few have shown effective transfer of training. We believe that, by softening the real/virtual divide, mixed reality technology has the potential to ease the transfer of rehabilitation activities into everyday life. We present two mixed reality systems, designed to support rehabilitation of activities of daily living and providing different mixtures of digital and physical information. Functional testing of these systems is described. System development and user evaluation continues, some of which is described in a sister paper (Edmans et al 2004) in this volume.

Virtual reality rehabilitation for all: Vivid GX versus Sony PlayStation II EyeToy, **D Rand**, **R Kizony** and **P L Weiss**, University of Haifa/Beit-Rivka Geriatric Medical Center, Petach-Tikva, ISRAEL

The main objective of this paper was to investigate the potential of the Sony PlayStation II EyeToy (www.EyeToy.com) for use in during the rehabilitation of elderly people with disabilities. This system is a projected, video-capture system which was developed as a gaming environment for children. As compared to other virtual reality systems such as VividGroup's Gesture Xtreme (GX) VR (www.vividgroup.com), the EyeToy is sold commercially at a relatively low cost. This paper presents three pilot studies which were carried out in order to provide essential information of the EyeToy's potential for use in rehabilitation. The first study included the testing of healthy, young adult participants (N=18) and compared their experiences using the EyeToy system to the GX system in terms of sense of presence, sense of enjoyment, control, success and perceived exertion. The second study assessed the usability of the EyeToy with healthy elderly subjects (N=10) and the third study assessed the use of the EyeToy with stroke patients (N=8). The implications of these three studies are discussed.

Memory assessment using graphics-based and panoramic video virtual environments, **A A Rizzo**, **L Pryor**, **R Matheis**, **M Schultheis**, **K Ghahremani** and **A Sey**, University of Southern California/Kessler Medical Rehabilitation Research & Education Corp., West Orange, NJ, USA

Virtual Reality (VR) technology offers new options for neuropsychological assessment and cognitive rehabilitation. If empirical studies demonstrate effectiveness, virtual environments (VEs) could be of considerable benefit to persons with cognitive and functional impairments due to traumatic brain injury, neurological disorders, learning disabilities and other forms of Central Nervous System (CNS) dysfunction. Testing and training scenarios that would be difficult, if not impossible, to deliver using conventional neuropsychological methods are now being developed that take advantage of the assets available with VR technology. These assets include the precise presentation and control of dynamic multi-sensory 3D stimulus environments, as well as advanced methods for recording behavioural responses. When combining these assets within the context of functionally relevant, ecologically valid VEs, a fundamental advancement emerges in how human cognition and functional behaviour can be assessed and rehabilitated. This paper focuses on the results of two studies that investigated memory performance in two VEs having varying levels of functional realism. Within these VEs, memory tests were designed to assess performance in a manner similar to the challenges that people experience in everyday functional environments. One VE used a graphics based simulation of an office to test object memory in persons with TBI and healthy controls and found that many TBI subjects performed as well as the control group. The other study compared healthy young persons on their memory for a news story delivered across three different display formats, two of which used a 360-Degree Panoramic Video environment. The results of this "in progress" study are discussed in the context of using highly realistic VEs for future functional memory assessment applications with persons having CNS dysfunction.

Preliminary evaluation of a virtual reality-based driving assessment test, **F D Rose**, **B M Brooks** and **A G Leadbetter**, University of East London, ENGLAND

Assessing one's own driving ability is very subjective, and there are occasions when an objective off-road assessment would be very useful, and potentially life-saving. For example, after physical or mental trauma, or approaching old age, it would be very useful for people to perform their own off-road assessment to help them to decide whether they should resume driving, or continue to drive. It is possible that people might be more likely to accept that it would be inadvisable for them to drive if they had themselves performed such an assessment. We are currently evaluating a virtual reality (VR) based driving assessment which runs on a PC and could be made easily accessible to people in these circumstances. The first stage of the evaluation was to evaluate the performance of drivers and non-drivers on the VR driving assessment and to compare the results obtained across the two groups of participants and with their performance on the Stroke Drivers Screening Assessment (SDSA). The VR driving assessment discriminated between drivers and non-drivers but the SDSA did not. In addition, two measures on the VR driving assessment correlated with drivers' scores on the SDSA.

Real-time clarification filter of a dysphonic speech and its evaluation by listening experiments, **H Sawada**, **N Takeuchi** and **A Hisada**, Kagawa University, JAPAN

This paper presents a digital filtering algorithm which clarifies dysphonic speech with the speaker's individuality preserved. The study deals with the clarification of oesophageal speech and the speech of patients with cerebral palsy, and the filtering ability is being evaluated by listening experiments. Over 20,000 patients are currently suffered from laryngeal cancer in Japan, and the only treatment for the terminal symptoms requires the removal of the larynx including vocal cords. The authors are developing a clarification filtering algorithm of oesophageal speech, and the primal algorithm of software clarification and its effectiveness was reported in the previous ICDVRAT. Several algorithms for the clarification have been newly developed and implemented, and are being evaluated by questionnaires. The algorithms were extended and applied for the clarification of the speech by the patients of cerebral palsy.

AudioBattleShip: blind learners cognition through sound, **J H Sánchez**, University of Chile, CHILE

Recent literature provides initial evidence that sound can be used for cognitive development purposes in blind children. In this paper we present the design, development, and usability testing of AudioBattleShip, a sound-based interactive environment for blind children. AudioBattleShip is an interactive version of the board Battleship game, providing different interfaces for both sighted and blind people. The interface is based on spatialized sound as a way of navigating and exploring through the environment. The application was developed upon a framework that supports the development of distributed heterogeneous applications by synchronizing only some common objects, thus allowing the easy development of interactive applications with very different interfaces. AudioBattleship was tested for cognitive tasks with blind children, evidencing that it can help to develop and rehearse abstract memory through spatial reference, spatial abstraction through concrete representations, haptic perception through constructing mental images of the virtual space, and cognitive integration of both spatial and haptic references.

AudioMath: blind children learning mathematics through audio, **J H Sánchez** and **H E Flores**, University of Chile, CHILE

Diverse studies using computer applications have been implemented to improve the learning of children with visual disabilities. A growing line of research uses audio-based interactive interfaces to enhance learning and cognition in these children. The development of short-term memory and mathematics learning through virtual environments has not been emphasized in these studies. This work presents the design, development, and usability of AudioMath, an interactive virtual environment based on audio to develop and use short-term memory, and to assist mathematics learning of children with visual disabilities. AudioMath was developed by and for blind children. They participated in the design and usability tested the software during and after implementation. Our results evidenced that sound can be a powerful interface to develop and enhance memory and mathematics learning in blind children.

Design and evaluation of a flexible travel training environment for use in a supported employment setting, **N Shopland**, **J Lewis**, **D J Brown** and **K Dattani-Pitt**, Nottingham Trent University/Learning Disability Services, London Borough of Sutton, Wallington, UK

This article describes the user centred design and development of a virtual environment (VE) to support the training of people with learning disabilities to travel independently. Three separate implementations were built on top of an initial design. Two of these environments implemented intelligent agents to scaffold learners using virtual environments; the third took stakeholder experiences to redesign the initial environment in an attempt to improve its utility.

Interactive rehabilitation software for treating patients with aphasia, **C Sik Lányi**, **E Bacsa**, **R Mátrai**, **Z Kosztyán** and **I Pataky**, University of Veszprém/National Centre of Brain Vein Diseases OPNI, Budapest, HUNGARY

Aphasia is an impairment of language, affecting the production or comprehension of speech and the ability to read or write. Most common cause of aphasia is – about 23–40 % of stroke survivors - acquired aphasia. The rehabilitation of aphasia is a medical, special treatment (speech therapy), which is the task of a psychologist. It needs long and intensive therapy. More detailed information about therapy can be found in (Engl at al, 1990, Subosits, 1986). In this paper we present our implementation or realization of interactive multimedia educational software to develop readiness of speech for helping the therapy. The software were developed within the frame of youth scientific and MSc thesis works. The first program was developed in Flash, the second in Macromedia Director. The goal of our software is to teach the most important everyday words. The software will be a useful device in the education of children with heavy mental deficiencies. Reading the program you can learn how it works and what current results we have achieved.

Using virtual public transport for treating phobias, **C Sik Lanyi**, **V Simon**, **L Simon** and **V Laky**, University of Veszprem/Semmelweis Medical University, Budapest, HUNGARY

Nowadays using and talking about virtual reality (VR) is a very popular subject. VR is an artificial world, a computer mediated environment. The user tries to enter fully into the spirit of her or his role in this unreal-world. Virtual Environment (VE) technology has undergone a transition in the past few years that has taken it out of the realm of expensive toy and into that of functional technology. During the past decade, in the field of Mental Healthcare, the considerable potential of VEs has been recognised for the scientific study. This paper shows the application of VR and presents the VR research in the University of Veszprem. The virtual worlds, introduced below, are developed for treating specific phobias (fear of travelling).

Problems with control devices experienced by people with intellectual disabilities using virtual environments: a systematic evaluation, **P J Standen**, **D J Brown**, **N Anderton** and **S Battersby**, University of Nottingham/Nottingham Trent University, UK

Virtual environments have a role to play in facilitating the acquisition of living skills in people with intellectual disabilities, improving their cognitive skills and providing them with entertainment. However, the currently recommended devices to allow navigation in and interaction with the environments are difficult to use. Using a methodology established in an earlier study, the study aims to systematically document the performance of users with the currently recommended devices in order to i) inform the design of a usable control device or devices and ii) act as a baseline against which they can be evaluated. 40 people with severe intellectual disabilities aged between 21 and 67 years used four environments with an equal number of sessions with the different devices being evaluated. Results suggest that for navigation, the joystick is better than the keyboard but that for interaction the mouse is better than using the fire button on the joystick. Preventing slippage of the joystick base would make its use much easier and it is suggested that separate devices are retained for navigation and interaction.

Providing external memory aids in haptic visualisations for blind computer users, **S A Wall** and **S Brewster**, University of Glasgow, UK

Haptic force feedback devices can be used to allow visually impaired computer users to explore visualisations of numerical data using their sense of touch. However, exploration can often be time consuming and laborious due to the "point interaction" nature of most force feedback devices, which constrains interaction to the tip of a probe used to explore the haptic virtual environment. When exploring large or complex visualisations, this can place considerable demands on short term memory usage. In this respect, a fundamental problem faced by blind users is that there is no way to mark points of interest or to harness external memory, in a similar way in which a sighted person may mark a graph or table at a point of interest, or leave a note in a margin. This paper describes the design, implementation and evaluation of external memory aids for exploring haptic graphs. The memory aids are "beacons" which can be used to mark, and subsequently return to, a point of interest on the graph. Qualitative evaluation by visually impaired users showed that external memory aids are a potentially useful tool. The most commonly reported problem was that of using the keyboard to control placing of the beacons. Suggestions for subsequent re-design of the beacons in light of the participants' comments are considered.

Game accessibility case study: Terraformers – a real-time 3D graphic game, **T Westin**, Stockholm University, SWEDEN

Terraformers is the result of three years of practical research in developing a real-time 3D graphic game accessible for blind and low vision gamers as well as full sighted gamers. This presentation focus on the sound interface and how it relates to the 3D graphic world, and also include post mortem survey results from gamers and comments to those.

Oxford

The City of Dreaming Spires

Rachel McCrindle

University of Reading

Welcome to Oxford, the city of dreaming spires, its skyline shaped by the golden stone buildings of the University with their spires, towers and domes. For over 900 years, Oxford has been a home to royalty and scholars, and since the 9th century an established town.

Although there is limited evidence of early Roman occupation around the city, it was really the Saxon Princess, St Frideswide, who around 727 AD founded a religious abbey on the site of Oxford's cathedral where Christ Church stands today. St Fideswide is often thought of as the 'founder' of the city, and it was the *Oxenford*, the original Saxon crossing point, that gave the city its name.

By the tenth century Oxford had become of strategic importance. Alfred the Great developed and fortified the city against the invading Danes and the city began to develop around the crossroads (Carfax) which is still the heart of the present city. After the Norman Conquest of 1066, King William gave Oxford to Robert D'Oilly who built Oxford's castle in 1072 and strengthened the city's ramparts. Richard the Lion heart was born at Beaumont Palace in 1157, and Charles 1 held parliament in the Convocation Hall during the Civil War.

Oxford University dates back to the 12th century and is the oldest English-speaking University in the world. Today it is made up of 39 colleges, each with its own governing body. There is no clear date of foundation, but teaching existed at Oxford in some form in 1096 and developed rapidly from 1167, when Henry II banned English students from attending the University of Paris.

In 1188, the historian, Gerald of Wales, gave a public reading to the Oxford dons and in 1190 Emo of Friesland became the first known overseas student. By 1201, the University was headed by a *magister scolarum Oxonie*, on whom the title of Chancellor was conferred in 1214, and in 1231 the masters were recognized as a *universitas* or corporation. In the 13th century, rioting between the townspeople and the students resulted in the establishment of primitive halls of residence which were later succeeded by the Oxford's colleges with University, Balliol and Merton Colleges, established between 1249 and 1264, being the oldest.

Oxford was soon recognized as being an eminent seat of learning. In 1355, Edward III paid tribute to the University for its invaluable contribution to learning and to the services rendered to the state by distinguished Oxford graduates. From 1878, academic halls were established for women, who became members of the University in 1920. Since 1974, all but one of Oxford's 39 colleges have changed their statutes to admit both men and women. St Hilda's remains the only women's college.

Early on, Oxford became a centre for research and lively debate, often of controversial issues, and it has continued in this vein as an international focus for learning and a forum for intellectual debate.

New College, our venue for ICDVRAT 2004, was founded in 1379 by the Bishop of Winchester, William of Wykeham (ca 1320-1404), about 200 years after the University came into existence. It is one of the largest and best-known colleges, and undeniably one of the most beautiful. The origin of the name 'New' stems from the fact that the College's 'official' name – the College of St Mary – is the same as that of Oriel College. Hence 'the new college of St Mary'. When founded, the College was new in several other ways. It was the first college for undergraduates, the first in which the senior members of the college had tutorial responsibility for undergraduates, and the first to be designed around the familiar quadrangle plan. Today, there are about 400 undergraduate and nearly 200 graduate students at the College.

Many leading luminaries have connections with Oxford and its University. Among these are 4 kings, 46 Nobel prizewinners, 25 UK Prime Ministers, 3 saints, 86 Archbishops and 18 Cardinals as well as many politicians, scientists, writers, broadcasters, actors and athletes. Recently celebrated has been the 50th anniversary of the first sub 4 minute mile, achieved by Sir Roger Bannister on the Iffley Road running track on 6th May 1954.

Oxford appears frequently as a setting for TV dramas and films such as *Inspector Morse*, *Iris*, *Shadowlands*, *The Saint*, *True Blue* and *Harry Potter*. It also has a long and distinguished literary history with many writers having spent time in Oxford and been inspired by the city. Amongst these authors are Lewis Carroll, JRR Tolkein, Oscar Wilde and Evelyn Waugh who based his famous novel *Brideshead Revisited* on his time at Hertford College.

It was during his time teaching at Christ Church that Charles Dodgson, a shy mathematics don, wrote the *Alice in Wonderland* stories under the pen name of Lewis Carroll. Alice was the daughter of Dean Liddell and the stories began one summer afternoon when Dodgson took Alice and her sisters out on the river. Alice's Shop in *Through the Looking Glass* was based on the shop opposite Christ Church in St Aldate's where the real-life Alice Liddell used to buy her barley sugar sweets.

Other notable children's favourites include *The Chronicles of Narnia* by C S Lewis and *The Hobbit* and *The Lord of the Rings* by J R R Tolkien. Both Lewis and Tolkien were leading members of The Inklings, Oxford's most famous literary group, that met regularly at The Eagle and Child pub in St Giles during the 1930's and 40's, to read extracts of their books to each other.

The Oxford University Museum of Natural History, where the conference reception will be held, houses the University's scientific collections of zoological, entomological, geological, palaeontological and mineralogical specimens, accumulated in the course of the last three centuries. The exhibits occupy a large centre court with its elegant cast iron columns supporting the great glass roof, and surrounded on four sides by upper and lower galleries. The museum is particularly renowned for it permanent display of dinosaurs, the largest display outside London; its reconstructions of the dodo, inspiration for the dodo in Lewis Carroll's Alice in Wonderland; and its live broadcasts to visitors of the museum of the swifts nesting in its tower. This study began in 1947 and is one of the longest running studies of any species of birds.

There are many other places of historical interest. These include Oxford's oldest building, the Saxon tower of St Michael at the Northgate, Cornmarket; the Bodlian library established over 400 years ago by Sir Thomas Bodley to house his collection of books; Sir Christopher Wren's Sheldonian Theatre built in 1664; the Ashmolean Museum, Britain's oldest museum and housing the university's world famous collections of art and archaeology; Christ Church whose college chapel is Oxford's cathedral, the smallest cathedral in England; the riverside botanic gardens opposite Magdalene College, and the garage in Longwall Street where in 1913 William Morris produced the first 'Bullnose' Morris Oxford thereby establishing Oxford as a leading centre of the motor industry.

Today, the city is a bustling town with a cosmopolitan range of restaurants and lively pubs, a wide variety of entertainment, excellent shopping and world-famous museums. Alongside the traditions of the University, Oxford is now also home to a growing hi-tech industry.

We hope that amidst the busy conference programme you will be able to find some time to explore all that Oxford has to offer. One good place to start might be the Oxford Story in Broad Street which, through a combination of films, interactive exhibits and a 25 minute indoor 'dark' ride, complete with the sounds and smells of the times, takes you through the University's 900 year history, allowing you to meet along the way, some of the writers, scientists and politicians whose careers began at Oxford.

ICDVRAT 2004

Session I. Stroke Rehabilitation

Chair: Rachel McCrindle

Mixed reality environments in stroke rehabilitation: development as rehabilitation tools

J A Edmans[1], J Gladman[1], M Walker[1], A Sunderland[2], A Porter[2] and D Stanton Fraser[3]

[1]Division of Rehabilitation and Ageing, Queens Medical Centre,
University of Nottingham, Nottingham, UK

[2]School of Psychology, University of Nottingham,
Nottingham, UK

[3]School of Psychology, University of Bath,
Bath, UK

judi.edmans@nottingham.ac.uk, john.gladman@nottingham.ac.uk, marion.walker@nottingham.ac.uk, alan.sunderland@nottingham.ac.uk, amandajayneporter@hotmail.com, d.stantonfraser@bath.ac.uk

ABSTRACT

A virtual or mixed reality environment for neurological rehabilitation should simulate the rehabilitation of the task, and not simply simulate the task. This involves identifying the errors in task performance that neurologically damaged patients make during task performance and replicating the guidance given during skilled rehabilitation. Involvement of a skilled therapist in the design and development team is essential. Neurological rehabilitation is complex and replicating it requires compromises between the desire to replicate this complex process and the amount of development time. Virtual or mixed reality systems that can simulate the rehabilitation process are suitable for clinical effectiveness studies.

1. INTRODUCTION

Loss of the ability to make a hot drink after a stroke is common. It is important because this loss makes returning home independently from hospital more difficult. Accordingly, this ability is commonly assessed during stroke rehabilitation in hospital and treatment aimed at restoring function is commonly given. We have developed virtual and mixed reality environments to aid the rehabilitation of this task.

1.1 Prior development of mixed reality environment for making a hot drink

A user-centred design process was used. Initial work, in which patients and Occupational Therapists were consulted, confirmed that virtual environment based systems could potentially be useful for this purpose (Hilton et al, 2000). Early experience with patient groups showed that acceptable means of interfacing with a virtual environment were needed before the virtual environment itself could be developed. A keyboard, mouse, or joystick was not appropriate for this group of patients. One approach, taken from earlier experience in the use of virtual environments with people with learning disabilities, was to develop a Tangible User Interface in which objects were manipulated to control the virtual environment instead of a mouse / joystick or keyboard (Hilton et al, 2002). Initially a keyboard-mounted device was developed in which movement of the objects simply generated pressed keys on the keyboard. The Tangible User Interface was found to be too inflexible. Consequently, a second approach used objects containing movement sensitive switches, thereby replacing the keyboard entirely. Objects in the first movement sensitive objects interface version connected to the virtual environment using cables, but in a later version the wires were replaced with radio transmitter devices. This provided a very flexible system, free of cables, keyboards, joysticks or mice.

Another approach to interface with the virtual environment has been to use an off-the-shelf touch screen interface, requiring no new technical development. Yet another approach has been to use machine vision of normal kitchen objects to drive the making of a hot drink in the virtual environment: this has required considerable technical development (Ghali et al, 2003). The development of three interface approaches (touch screen, movement sensitive objects interface, machine vision) has enabled us to develop a mixed reality system, which offers the opportunity to train both the cognitive and physical aspects of task performance. Details and justification of this approach are given in another paper (Hilton et al 2004).

Once workable interfaces had been developed, it was then possible and necessary to develop our VIRTOOLS virtual kitchen environment controlled by these interfaces. At this stage the elements of an acceptable and usable system for people with stroke had been developed.

The development team until this stage included experienced technical staff, and had consulted with health care staff including Occupational Therapists, but did not include an Occupational Therapist within the team, and no field tests with stroke patients undergoing rehabilitation had been undertaken.

In this paper we report our experience during the necessary work to develop our system for making a virtual hot drink into one that was not only usable by people with stroke, but was also fit for purpose as a rehabilitation tool.

1.2 Virtual environments as rehabilitation tools

The development work so far had focussed extensively and necessarily on means to interface with the virtual environment. In this next phase we had to concentrate more upon the virtual environment itself and how the system would be used as a rehabilitation tool in clinical practice. This required attention to the reasons why these types of technologies have potential value in rehabilitation, and why they may fail. The potential rehabilitation benefits of practice in a virtual environment include:

- Treatment of the cognitive processes of task performance can take place sooner in a virtual environment than in a real environment. For example, patients with physical problems might be able to rehearse the cognitive aspects of task performance needing only sufficient motor capacity to control a virtual environment rather than real objects.
- Virtual environments can avoid retraining in potentially hazardous settings. For example, patients would not be at risk from boiling water or electricity as they would be if they were in a real environment.
- Working with virtual environments can be enjoyable and compulsive, so providing motivational benefits. Many rehabilitation experts believe that motivation during rehabilitation is a factor in its success. One reason for this is simply that patients who are more motivated spend more time and effort into promoting their recovery and that this drives neural plasticity, the presumed underlying mechanism of neurological recovery after brain injury (Taub et al, 2002). Virtual environments can be designed so that patients can use them with little or no clinical supervision. This could increase the amount of time spent in rehabilitation.
- Virtual environments can allow the learning process to be more strictly controlled and defined than might be possible in the real environment. “Scaffolded” learning could be used (Wood et al, 1976), in which the task is initially simplified and then progressively made more complex as the patient’s ability increases. This process is called “shaping” in the neurorehabilitation literature (Taub et al, 2002). The distractions of real world learning, and de-motivation caused by repeated failure if too complex a task is attempted, can be avoided.

The potential rehabilitation drawbacks of virtual environments are:

- Interfacing with the virtual environment could be too difficult for patients to use. Patients cannot benefit from rehabilitation if they do not participate in it.
- The system could be too difficult, unrealistic or intuitive to be enjoyable and motivational. Patients may not benefit if not motivated.
- The system could fail to train the cognitive skills that are lost. Brain injuries such as stroke cause a huge variety of neurological impairments, singly or in combination, and these differ between individuals. It is important to design environments and training routines that deal with the common faults and errors neurological patients make during making a hot drink, which may differ from those made by non-neurologically impaired people.
- Training in the virtual environment could fail to generalise to real world settings. It has been found that people with cognitive problems due to acquired brain damage such as stroke often show poor generalisation from a training task to performance in naturalistic settings (Manly, 2002). Little is known about the benefits of mixed reality rehabilitation environments.

2. DEVELOPMENT OF THE SYSTEM AS A REHABILITATION TOOL

Further to the development of a system that was usable by stroke patients, where a virtual hot drink could be made, we now needed to develop the system so that it simulated rehabilitation of this task. In particular we needed to ensure that the learning processes invoked during use of our system were likely to promote recovery. For example, an early version of our system used in usability trials simply instructed the user through the steps of making a hot drink. Clearly, that version did not provide scaffolded learning, or opportunities for problem solving. In this stage therefore, we required the typical errors in performance made by patients to be identified, and the presumed effective parts of the guidance and feedback used during rehabilitation to be replicated. For this, we needed the contribution of an Occupational Therapist experienced in stroke rehabilitation. The following observations were made.

2.1 Task model

The Occupational Therapist observed that many stroke patients do not drink or want to make coffee (the task within the virtual environment at this stage) and therefore they would not see this as a meaningful part of their rehabilitation. Both tea and coffee making therefore needed to be included in the task model and the virtual environment.

The Occupational Therapist also observed that there is no set procedure for making a hot drink and people do this task in many different ways, e.g. some people put milk in the mug before the tea and some do it after the tea, some people even put everything into the kettle and boil them up together. In clinical practice the therapist would therefore find it inappropriate to try to train a pre-determined, "correct" method for making a hot drink. Instead, the therapist would teach the patient to do the task the way they want to do it, providing it is functional and safe. The task model therefore had to be examined and adjusted so that several means of completing it were permissible, not one alone.

2.2 Virtual environment

In clinical practice, the Occupational Therapist often attempts to simulate the patient's home setting including the range of objects, colour of objects, shape of objects, position of objects, other clutter, etc. The ability to do this in a hospital or clinic setting is limited, but is potentially possible in a virtual environment. However to do so, would need the system to be reconfigured each time, and would require a considerable further development. The need to reconfigure the system would put increased demands upon the Occupational Therapist. By contrast, a simple model that does not need configuration is likely to be more widely applicable.

In clinical practice, milk jugs, mugs and kettles can be almost any colour and it would not have been difficult to allow different shaped and coloured objects to be selected in the virtual environment. However a constraint of our mixed reality system is that the machine vision approach we used involves recognition of colour, rather than shape, to recognize objects. This meant that a different colour had to be used for each object (i.e. green kettle, blue mug, etc) and these could not be modified easily. Thus the use of machine vision makes configuration more complicated.

It is possible that a lack of configurability may mean that patients find the virtual environment too unreal or perceive it to be a game or toy, with the implication that it is a frivolous activity. If so, they may not be motivated to engage in it. However, our experience has been that patients do not need perfect simulation. For example, in our system, a "ping" sound with a white arrow flashing above the object was used to denote that a virtual object had been selected, and when the virtual environment gives a prompt to move an object, the object "flashes" indicating which object needs to be moved. These unreal effects did not seem to trouble either the patients or the Occupational Therapist.

We elected to keep the model simple, on the basis that we did not feel that additional complexity was necessary at this stage, but also because of limitations upon development time.

2.3 Identification of errors

People with stroke have many different and complex impairments of varying severity, many of which reduce their capacity to make a hot drink. These may include weakness of the arm, impaired vision, a variety of cognitive and executive problems or a combination of these impairments. In clinical practice, errors during attempted task performance are noted by therapists and used to infer the likely underlying impairments. The Occupational Therapist then helps the patient to find means to overcome these problems. Helping the patient to overcome these errors during task performance is the rationale for occupational therapy: promoting recovery through purposeful activity.

The first challenge this poses in the development of the virtual environment as a rehabilitation tool is that, like a therapist, it must be able to detect when a pathological error (i.e. an error due to the stroke, as opposed to a simple error caused by carelessness or unfamiliarity with the system) has occurred.

To deal with this problem, the Occupational Therapist repeated a task analysis for making a hot drink, with the intention of creating a list of all the possible stages required to make a hot drink. Twenty-five stages were identified. Some of the stages were essential, such as getting the kettle, and other stages were optional, such as spooning sugar into the mug. Using the touch screen interface to the virtual environment involves 21 of the stages. Use of the movement sensitive objects interface alone only involves 10 of the stages, although all are involved when the movement sensitive objects and vision systems are used as a single interface (Hilton et al, 2004).

Next, the Occupational Therapist also compiled a list of the different types of pathological error types frequently made by patients when making a hot drink. These error types were categorised empirically into four groups: attention; sequencing; object use; dexterity and accuracy. Occupational Therapists identify these pathological error types from an almost infinite number of incorrect actions that stroke patients may make whilst making a hot drink. These include leaving the kettle lid up whilst boiling (it will not switch off with the lid up), holding a mug to the kettle spout instead of tipping the kettle to the mug, or putting insufficient water to cover the element in the kettle. These incorrect actions could be due to a variety of error types. For instance, if a patient left the kettle lid up whilst boiling, the therapist would try to ascertain whether this was due to an attention error or an object use error. Each patient's rehabilitation is then based on their impairments, the types of incorrect actions made and the possible types of errors.

Having established the types of errors commonly seen during stroke rehabilitation, we then reviewed which interface approach (touch screen, movement sensitive objects interface/machine vision) might be suitable for detecting incorrect actions and for treating patients with the different types of error. Our findings are presented in the table below.

Our interpretation of this analysis is that our mixed reality system would be able to detect most of the errors made by stroke patients undergoing rehabilitation. As expected, the movement sensitive objects / machine vision interface would be able to detect errors in the physical aspects of task performance. Although we have not had the time to do so, it would be possible to develop our system so that the degree and duration of tilt of the kettle would affect the speed the mug is filled with boiling water, and allow it to overflow. Similar changes could be made to the filling of a water jug, the kettle and so on.

We did not feel that our system could be easily used without an Occupational Therapist to interpret incorrect actions. This limits the extent to which our system can be used unsupervised and, hence, potentially limits the opportunity for increasing the amount of time a patient spends in rehabilitation activity.

2.4 Prompts and guidance

The second challenge to the development of the virtual environment as a rehabilitation tool, after identification of pathological errors, was to replicate the effective aspects of prompts and guidance used during rehabilitation.

To do this, the Occupational Therapist videoed several clinical sessions during which she assessed hot drink making ability in patients in a real kitchen and when attempting to do so using our mixed reality system. Together with a colleague, she identified that she used a three-stage process. First, a prompting question would be used such as "What would you do now?" If this did not produce the correct behaviour a more direct instruction would follow, such as "Put the teabag into the teapot". Third, if this did not produce the desired action, a visual demonstration would be given. These prompts were individualised to suit individual patients' impairments.

The initial questions were felt to be the most important part of the educative process because they tended to prompt problem solving rather than provide ready-made solutions. Examples include: "What do you need now?" or "What do you need to do next?" These questions were used to concentrate on problems of sequencing, attention and initiation, when the patent was stuck, when the error was failure to proceed to the next step. Another question type was to ask patients to reflect on what stage they have just completed and ask a question. For example "You have cold water in the kettle, will it boil like that?" This can be used to question object use, sequencing, problem solving or safety issues. These questions were easily incorporated into the mixed reality system in response to errors in performance of the virtual task.

However, not infrequently the therapist had to ask the patient "What are you trying to do?" in order to understand the patient's behaviour or to identify the type of error (for example, if the patient had difficulty locating an object). In practice, this enquiry was not simple, and often involves a conversation between the therapist and patient, including the interpretation of non-verbal information. We did not think this part of the process was likely to be easily achieved through interaction with the system alone. For that reason we felt that our mixed reality system alone is unlikely to be useful without the Occupational Therapist to supervise.

Table 1. *Categories of errors and suitability of interfaces for their rehabilitation*

TYPE OF ERROR		SUITABILITY OF INTERFACES
ATTENTION		
Initiation	Does not automatically begin the task or a stage	Detectable using all interfaces
Attention	Does not attend to an individual event (e.g. pays no attention to the fact the kettle has boiled)	Detectable using all interfaces
Neglect	Cannot find an object or does not respond to a visual or auditory cue to the affected side (e.g. unable to locate teapot positioned on his/her affected side)	Detectable using all interfaces but will need an OT to interpret behaviour
SEQUENCING		
Sequence	Performs an action at the wrong time within the task (e.g. puts the sugar in the cup before the teabag)	Detectable using all three
Addition	Adds an abnormal action (e.g. rips a teabag open and pours loose tea into the teapot)	Will need an OT
Omission	Omits a stage (e.g. fails to put any water in the kettle)	Detectable using all interfaces
Perseveration	Repeats a stage (e.g. pours the milk into the cup twice)	Detectable using all interfaces
OBJECT USE		
Selection	Does not select the correct object to accomplish a stage (e.g. stirs the tea with a finger or pours milk into the teapot)	Detectable using all interfaces but will need an OT to interpret behaviour
Object Use	Does not use object appropriately (e.g. uses the kettle as a teapot)	Detectable using all interfaces but will need an OT to interpret behaviour
Problem Solving	Gives unmistakable signs of not knowing what to do (e.g. continues to place the cup near the spout of kettle without picking up the kettle or looks hesitatingly at the objects, picking them up, turning them over, putting them down and trying with another object.)	Detectable using all interfaces
DEXTERITY & ACCURACY		
Dexterity	Fumbles when attempting to use objects (e.g. spills coffee when spooning)	Best detectable using movement sensitive objects and machine vision interfaces
Quantity	Misjudges the amount of something (e.g. fills the cup with more milk than tea)	Not currently detectable
Spatial	Misjudges the location of objects (e.g. misses the cup and pours the tea onto the table)	Not detectable using movement sensitive objects interface alone.

2.5 Feasibility

The original aims were for the virtual environment to be suitable for use at the patient's bedside in a hospital ward situation and to be safe for use by both patients and staff. Consideration therefore needed to be given to the security and portability of the virtual environment in this setting, any disturbance it may cause, plus the confidentiality of patient information stored in the virtual environment.

The virtual environment currently includes:

- Laptop with touch screen interface
- Second laptop connected to a tangible interface and visual monitoring interface
- Computer table with a video camera mounted on a bracket above the table
- Tangible interface i.e. a set of kitchen objects with sensors attached

The amount of space and equipment involved presented a problem for making the virtual environment portable. In its present state, it takes a long time to set up and requires two members of staff to do so. The long-term plan is to build a suitable workstation on wheels, so the virtual environment can be left set up ready to use. This will resolve the portability problem but then requires a hospital ward to have sufficient space to store the virtual environment.

The equipment presents a security risk: theft is common in hospitals, and computers are particularly vulnerable. There is also a need to protect confidential patient information that might be stored in it. Therefore, the virtual environment cannot be left unattended, for instance, whilst not in use or even whilst taking a patient to the lavatory during treatment.

The majority of patients in our hospital are in bays of 4-6 beds and only a minority are in individual side rooms. This presented problems for using the equipment by the patient's bedside, due to the disturbance to other patients and staff in the bay, caused by the verbal prompts and sound effects from the virtual environment. To resolve this, the Occupational Therapist had to take the patient and the equipment to a more suitable room, such as the rehabilitation room, dayroom or single room. This is another factor that makes our system more suitable as an occupational therapy aid, rather than a system for independent use by the patient.

3. FIELD TESTS WITH STROKE PATIENTS

To assess the clinical relevance of the mixed reality environment as a rehabilitation tool, field tests were conducted with stroke patients. These field tests so far have been conducted primarily using the touch screen interface with 10 patients. The tangible / machine vision interface field tests are in progress.

Despite the large amount of preparatory work on making the virtual environment system accessible to stroke patients, the Occupational Therapist found that she needed to train patients in the use of the touch screen, so that they knew when they had selected or moved a virtual object. For this reason, a brief training programme, involving putting a letter in an envelope and sticking on a stamp was developed.

Commonly we found that patients spoke to the system as if it were a person. We found this encouraging as it indicated that patients were activity engaging in it. They found the system challenging, and generally indicated this with good natured comments indicating mild frustration when they made errors. Examples include "*She doesn't trust me does she? Don't blame her*" (error prompts are given by a female voice): to the prompt "*What would you do next*?" the reply was "*Throw it out the window*". Only one patient became quite angry due to frustration. However, despite these frustrations, most patients after using the system were complimentary about the system and its purpose, including the person who became angry. The system was described as *"Clever", "Good for people who find this difficult" and "Easier, cleaner and quicker that the real task".*

However, the field tests have indicated that the touch screen interface is probably not suitable for all patients. Those with limited vision had difficulty seeing the objects clearly on the screen. Patients with poor upper limb co-ordination had difficulty with accuracy when using the touch screen pen to operate the virtual environment. Patients with poor short term memory had difficulty remembering how to use the touch screen pen. Patients with poor hearing had difficulty hearing the instructions given the virtual environment.

Some patients were not comfortable with feedback coming from the virtual environment rather than the Occupational Therapist. Feedback from the virtual environment is at a set speed and some patients might want it quicker, or slower. For example, to the question "*What do you need to do next*?" one patient said, irritably, "*I'm about to tell you".*

Many users (both staff and patients) had difficulty relating the virtual task to the real world task. For example some tried to put the coffee jar into the mug without taking the lid off the jar and putting the spoon into the coffee jar. They said that it was in some respects harder to do the task in the virtual environment than in the real world.

4. FURTHER EVALUATION

The experiences reported here, including an Occupational Therapist in the design and development team, have improved the degree to which our rehabilitation system simulates the rehabilitation of the task of making a hot drink in stroke patients. We have developed the task model to make it reflect the variety of ways patients make hot drinks. We have checked that our system is able to identify the typical errors made by stroke patients during task performance, and we have developed a system of prompts that mimic the problem-solving process used in clinical rehabilitation practice. It seems to be well received by the sorts of patients for whom it is intended. In these respects, we have developed a system that is fit for purpose as a rehabilitation tool.

On the other hand, we recognise the limitations of our system. It needs to be used by an Occupational Therapist rather than to replace her. Some errors in real world performance may not be detected when performing the task using the mixed reality system. Inflexibility of the system, for example in terms of configurability, may hamper active involvement. The cognitive demands necessary to use the system may prove a barrier to it is use in some patients, as may the effect of poor vision or poor hearing.

Some of the limitations of our system could be overcome with yet more development of the virtual environment. However, ultimately the test of effectiveness of a rehabilitation tool is not simply that it can be used, or the degree of similarity between the virtual and real environments. We feel that the most important question now is whether there is any value in using our mixed reality environment in clinical practice. We are now conducting a randomised controlled trial, using the system in patients undergoing in-patient rehabilitation after stroke.

5. CONCLUSIONS

This stage of development of a mixed reality environment system took a system to make a virtual hot drink, with interfaces that made it usable by stroke patients, and developed its ability to simulate the rehabilitation of that task. The process required an Occupational Therapist who was experienced in stroke rehabilitation. This work required us to review the underlying task model for making a hot drink, to undertake an analysis of the errors made during neurological rehabilitation, and an analysis of the sorts of prompts and guidance given during skilled stroke rehabilitation. A system has been developed that is fit for purpose and, although further development could continue, it now requires further evaluation using clinical effectiveness studies.

Acknowledgements: We would like to acknowledge the collaboration of the rest of project team at the University of Nottingham: Dr Sue Cobb and Dr Richard Eastgate, of the Virtual Reality Applications Research Team (VIRART), Jonathan Green, Dave Hilton and Dr Tony Pridmore, of the School of Computer Science. We would also like to thank the Stroke Association for funding the evaluation of this environment.

6. REFERENCES

A Ghali, A Cunningham and T Pridmore (2003), Object and event recognition for stroke rehabilitation, *Proceedings of Visual Communication and Image Processing Conference, 2003.* Details available on *http://www.cs.nott.ac.uk/IPI/publication29.html*

D Hilton, S V G Cobb and T Pridmore (2000), Virtual reality and stroke assessment: Therapists' perspectives, *Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies 2000,* Alghero, Italy.

D Hilton, S Cobb, T Pridmore and J Gladman (2002), Virtual reality and stroke rehabilitation: A tangible interface to an every day task, *Proceedings of the 4th International Conference on Disability, Virtual Reality and Associated Technologies 2002,* Veszprem, Hungary.

D Hilton, J Green, T Pridmore, R Eastgate and S Cobb (2004), Mixed reality environments in stroke rehabilitation: interfaces across the real / virtual divide, *Proceedings of the 5th International Conference on Disability, Virtual Reality and Associated Technologies 2004,* Oxford, UK.

T Manly (2002), Cognitive rehabilitation for unilateral neglect: Review, *Neuropsychological Rehabilitation*, **12**, 4, pp. 289-310.

E Taub, G Uswatte and T Elbert (2002), New treatments in neurorehabilitation founded on basic research, *Nature Reviews Neuroscience*, **3**, 3, pp. 228–236.

D Wood, J S Bruner & G Ross (1976), The role of tutoring in problem solving, *Journal of Child Psychology and Psychiatry*, **17**, 2, pp. 89-100.

Mixed reality environments in stroke rehabilitation: interfaces across the real/virtual divide

T Pridmore[1], D Hilton[2], J Green[1], R Eastgate[3] and S Cobb[3]

[1]School of Computer Science and IT,
University of Nottingham, UK

[2]School of Nursing,
University of Nottingham, UK

[3]VIRART
University of Nottingham, UK

Tony.Pridmore@nottingham.ac.uk

ABSTRACT

Previous studies have examined the use of virtual environments (VEs) for stroke and similar rehabilitation. To be of real benefit it is essential that skills (re-)learned within a VE transfer to corresponding real-world situations. Many tasks have been developed in VEs, but few have shown effective transfer of training. We believe that, by softening the real/virtual divide, mixed reality technology has the potential to ease the transfer of rehabilitation activities into everyday life. We present two mixed reality systems, designed to support rehabilitation of activities of daily living and providing different mixtures of digital and physical information. Functional testing of these systems is described. System development and user evaluation continues, some of which is described in a sister paper (Edmans et al 2004) in this volume.

1. INTRODUCTION

Stroke is a term used to describe a sudden neurological deficit within the brain. The extent and precise location of the damage is unique to the individual, making intact function and observed behaviour also individual to each stroke survivor. A thorough assessment of the patient's cognitive and motor function is the first stage of rehabilitation. For most patients the priority is to facilitate a return home as soon as it is safe and timely for them to do so.

Rehabilitation necessarily aims to restore function and so may be aimed at reducing impairments or promoting activities. Following a series of workshops and seminars with stroke survivors, occupational therapists and consultants, the activity of making a hot drink was selected as a suitable activity of daily living upon which to base the research reported here (Hilton et. al. 2000, 2002).

A number of previous studies have examined the use of virtual environments for stroke rehabilitation focussed on activities of daily living. Davies et al (2002) used virtual reality in three different scenarios: a kitchen activity, operating an ATM (cashpoint) and way finding. A meal preparation task in a virtual kitchen was the focus of research by Christiansen et al (1998), while Brown et al (1999) considered a variety of activities taking place within a virtual city. A virtual environment to train stroke survivors to cross a street safely was the focus of work by Weiss et al (2003). Gourlay et al (2000) have developed a hot drink task as a virtual environment for rehabilitating stroke survivors accessed via a mouse or data glove.

Virtual environments (VEs) have as their core the simulation by computer of three-dimensional space; they can be explored in real time with similar freedom to real world exploration, and the user may interact with objects and events in the simulation. Interactions with VEs reproduce similar visual-spatial characteristics to interactions with the real world, and they can preserve the link between motor actions and their perceived effects (Regian, Shebilske and Monk, 1992). This has lead to a focus on virtual rehabilitation environments, in which survivors of stroke and those with other, similar conditions rehearse tasks that would be problematic in the real world. It is often pointed out that rehabilitation tasks presented within a VE enable patients to repeat tasks in safety, to feel free to manipulate the world autonomously and, if the experience is an enjoyable one, gain much needed confidence. To be of any real benefit, however, it is essential that skills (re-)learned within a virtual environment transfer to corresponding real-world situations. While many tasks have been developed in VEs, only a limited number have demonstrated effective transfer of training (Linden et al. 2000; Mendozzi, Pugnetti, Barbieri et al 2000; Rose, Brooks & Attree, 2000; Stanton et al., 2000).

A potential alternative to the self-contained virtual rehabilitation/learning environment is provided by the recent development of mixed reality environments and systems. Mixed realities are spatial environments in which participants can interact with both physical (real) and digital (virtual) information in an integrated way (Milgram and Kishino 1994). Mixed reality technologies have been employed in a variety of entertainment, art and educational scenarios, but have yet to be explicitly and systematically applied to rehabilitation. By softening the real/virtual divide we believe that mixed reality technology has the potential to ease the transfer of rehabilitation activities into everyday life. This might be achieved either by making critical physical information available during a single and otherwise virtual rehabilitation activity, or by performing that activity in a series of increasingly physical mixed reality systems over an extended rehabilitation programme.

The adoption of a mixed reality approach naturally focuses attention on the technology used to interface the real and virtual environments. The provision of effective and usable interfaces to virtual rehabilitation environments is non-trivial; standard interface technologies are frequently inappropriate per se, and a wide variety of user abilities must be provided for. Previous work on interfaces to VEs, including our own (Cobb et. al. 2001), has, however, focussed on providing the user with access to or a sufficiently high level of immersion in a virtual rehabilitation environment. Residual perception of the surrounding real environment is seen as a shortcoming. Mixed reality work differs from traditional virtual reality in two key respects. First, emphasis is placed on the complete system, not just the virtual environment. Second, information regarding the real environment is viewed as a resource, not a problem. We adopt these views here.

2. A MIXED REALITY FRAMEWORK

Mixed reality technologies can be characterised by their relative positions along an axis spanning the real/virtual divide (Figure 1). Immersive virtual reality, experienced via head-mounted displays or CAVEs (Cruz-Neira et. al. 1992) provide the most completely virtual experience; participants can become involved in the virtual environment to the complete exclusion of the surrounding physical world. In augmented virtuality (Milgram and Kishino 1994) representations (e.g. images or video streams) of real objects are included in the virtual environment, allowing the inhabitant of a VE to access physical information. Mixed reality boundaries (Benford et al 1996) mark the midpoint of the continuum between real and virtual environments. Mixed reality boundaries connect virtual and physical spaces by creating a transparent boundary between them. Projective displays allow inhabitants of the real world to view events in the virtual, while co-located cameras or other sensor technology allow those in the virtual environment to view the physical. Mixed reality boundaries with a variety of properties are in existence (Koleva et. al. 1999). Moving further towards the physical, augmented reality overlays digital data on views of the real world, usually via a transparent display (e.g. Billinghurst et al 1996), allowing the user to view but not usually to manipulate digital information. The physical manipulation of digital information is, however, key to the notion of "tangible bits" proposed by Ishii and Ulmer (1997). Tangible bits uses graspable physical objects to manipulate digital data, so the movement of an object in the physical world has a corresponding and predictable effect on the virtual.

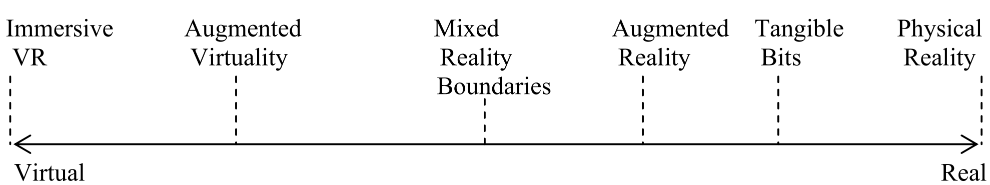


Figure 1. *Technologies across the real/virtual divide.*

The technologies overviewed above each provide the user with a unique mixture of the real and the virtual. This mixture should be exploited and explored to determine which provide(s) the most appropriate experience(s) at each stage in the rehabilitation process. With this goal in mind we have developed a variety of interfaces to a virtual environment designed to support the making of a hot drink. The resulting mixed reality systems can be characterised by their position on the continuum of Figure 1. Following a brief account of the common virtual environment they are described below.

3. THE VIRTUAL ENVIRONMENT

The role of the VE is to encapsulate and support the presentation of the digital information made available to the user (patient). Here this comprises a model of the task at hand, expressed via direct visual feedback and pre-recorded audio-visual guidance and demonstrations. A hierarchical task analysis (HTA) was developed to divide the top-level activity of making a hot drink into progressively smaller discrete subtasks, the process continuing until the subtasks represent individual actions.

Close integration of the VE and any sensors providing information from the physical world is critical, and for this reason we have shifted VE development from SuperScape™ to Virtools™, in recognition of the latter's enhanced ability to communicate with external software and devices. Objects were modelled in Lightwave ™ and are based on real kitchen objects used in patient assessments. The functional status of each object is maintained and used to determine which actions are permissible/desirable. Suitable words and phrases were compiled (Edmans et. al. 2004) and a health care of the elderly nurse with experience of multimedia recording provided the voice-over. Figure 2 shows a sample view of the environment.

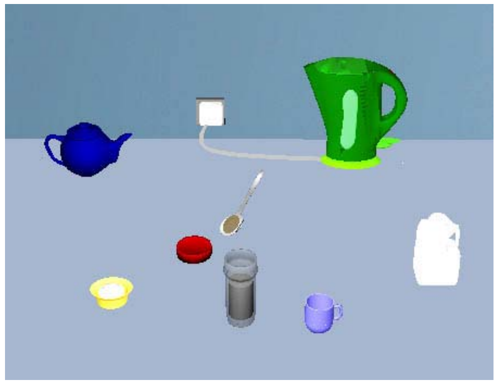


Figure 2. *The kitchen activity layout in Virtools™.*

In operation, each interaction is logged and scored according to a scheme devised in collaboration with an Occupational Therapist (Edmans et. al. 2004). A correct action scores 2 points. If after 20 seconds no input has occurred a verbal prompt is given and subsequent actions score one point. Following a further time period a verbal instruction is given, informing the user what to do. If there is still no response a demonstration of the correct action is provided and the user scores 0 for that subtask. To test the functionality of the scoring system 20 representative scenarios were designed and scored manually before being executed within the VE by one healthy user. In each case the automatically generated scores were as expected.

4. THE TOUCH SCREEN: A MIXED REALITY BOUNDARY

Touch-sensitive screens offer a direct mode of controlling objects in a virtual world. A touch screen is a simple example of an (asymmetric, Koleva et al 1999) mixed reality boundary, and so lies at the midpoint of the continuum of Figure 1. Recent developments have made the touch screen an affordable option that might offer an opportunity for patients immediately post-stroke to commence rehabilitation using a notebook computer from their hospital beds. Providing a single, localised view of the virtual environment, this technology may be of particular use to individuals with diminished ability to divert attention between multiple locations. Hofmann et. al. (2003) used touch screens to train Alzheimer's patients on a shopping task and attribute the positive effects of their training regime in part at least to the ease of use of the interface.

A lightweight and robust Toshiba Satellite Pro A10 notebook computer with a 14" touch sensitive monitor was selected and alternative touch screen controls developed. A two touch operation was devised to provide the simplest form of input: an action is be initiated by selecting appropriate objects or pairs of objects. Identification of successful selection is an important design feature and feedback has been incorporated in the form of a pointer accompanied by a brief audible tone. The base functionality of the system was tested by taking a typical scene from the hot drink environment, selecting all possible pairs of objects and ensuring that the required/expected model subtask was generated. A drag and drop mode was offered as an alternative, relying upon proximity detection between objects. While they may slow interaction, Potter et al (1988) show that more realistic but more complex interactions of this type are acceptable to touch screen users and can produce fewer selection errors.

To compare the touch screen controls, a simple task was implemented in Virtools™. This required the user to apply a clearly visible stamp to a clearly visible envelope. In two touch mode the user must first select any point on the stamp, followed by any point in the upper right quadrant of the envelope. The stamp then moves to a predefined position in the corner of the envelope. If the second point is not in the acceptable area an audible error message is generated. In drag and drop mode, selecting and dragging the stamp onto the upper right quadrant causes the stamp to fly to the predefined stamp position when the finger is released (the

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

“take-off” strategy of Potter et al 1988). If the stamp is released outside that quadrant it remains at the point at which it was released and an audible error message is generated.

The stamp task was attempted by nine patients, at various stages post-stroke, under the supervision of an experienced OT. Each made ten attempts using each of the touch screen controls. Every patient managed to successfully complete the task within ten trials using the two touch interface, but none could reliably use the drag and drop system. Several were forced to use non-dominant hands, but even those using dominant hands had difficulty. When asked which they would prefer to use for the hot drink task all selected the two touch version. Lack of accuracy at anything above a very slow speed was cited as the primary cause of difficulty, though the need to steer round other objects generated noticeable, if secondary, problems. In a study of brain-injured patients by Linden et al (2000), all subjects expressed a desire for a drag and drop mode. Our studies found that the precision required to complete a drag and drop action using a touch screen significantly increases the difficulty of the task for stroke patients.

Following a period of frequent design iteration to determine system function, 10 patients admitted to a stroke unit have been involved in a continuing programme of participatory design informing the development of the VE/touch screen system. Patients completed the hot drink activity both independently and with OT supervision. Video recordings, field notes, verbal feedback and automatic data recorded in the VE have been used to assess the developing system. Details are given in a sister paper (Edmans et al 2004). The only issue raised regarding touch screen control was the standard difficulty in selecting small objects. Observed sequences of object selection are consistent with the task model, indicating that identification of objects in the VE does not appear to have been an issue with these particular patients. Although common, requests for increased realism centred on realism of behaviour, not of appearance. Evaluation of the system continues.

5. INCREASING REALISM: A TANGIBLE INTERFACE

Ishii and Ulmer (2001) define tangible user interfaces (TUIs) as giving “physical form to digital information, employing physical artefacts as representations and controls of computational data”. As such they provide more physical information than do mixed reality boundaries. The use of TUIs is rare in rehabilitation, though Sharlin et al’s (2002) “Cognitive Cubes” have shown some potential as a cognitive assessment tool.

Cobb et al. (2001) describe TUIs to an everyday task designed for use by young adults with a learning disability. Studies with stroke survivors in the community criticised these early systems, in which the graspable objects were mounted on a base board, as lacking flexibility (Hilton 2002). The ability to reposition objects is considered to be essential, as is the freedom to complete the task in a preferred sequence. To improve mobility the base board was removed. Real kitchen artefacts were retro-fitted with compact sensors and machine vision techniques employed to recognise and recover the position of individual objects.

Non-mercury tilt switches were mounted inside vessels intended for pouring. Sub-miniature micro-switches were mounted inside ABS boxes with 49x29.5mm plastic lids mounted on rubber shims to provide pressure operated switching. A test bed was developed in which twin cables connected each sensor and an input pair of a keyboard encoder. A connection made across an input pair emulates a key press, allowing individual sensors to provide unique codes. In a single case study, part of the iterative design process, a 60 year old stroke survivor unable to use the left side of his body completed the hot drink making activity using the sensor driven TUI, demonstrating that the sensors operated as expected. A wireless system was then constructed using RF Solutions Ltd AM-RT4-433 Transmitters and AM-HRR3-433 Receiver operating at the general purpose telecommand / telemetry band of 433MHz frequency. The wireless system has been tested up to a distance of 2m from the receiver and found to be effective.

Object position is recovered from colour images provided by a vertically mounted digital camera whose field of view covers the workspace. As real time response and robustness are of primary importance we follow Swain and Ballard (1991) in basing object recognition on colour histograms. An 8*8*8 bin histogram of the colours present in each camera frame is compared with a set of pre-calculated model histograms, one per object. As objects can be chosen to be fairly colourful in this environment, each model histogram will contain peaks at certain chromaticities; the image histogram should contain similar peaks if the object is present. A given bin of the image histogram may, however, contain more pixels than the corresponding model bin. These extra pixels should be discounted. The difference between the value $N_I(i)$ in the i^{th} bin of the image I and the value $N_M(i)$ in the i^{th} bin of the model M is calculated. The minimum of zero and this calculated difference forms the basis of the histogram comparison operation. Summing this over the image I gives a value between 0 and $-N_M$ where N_M is the total number of pixels in the model histogram. If all bins in the image contain at least as many pixels as the corresponding model histogram bin, the result is 0, if the image histogram has no pixels in the bins in which the model histogram pixels are found – if the occupied sets of bins from the histograms do not intersect – then the result is $-N_M$. To produce a fractional estimate we use:

$$C(M,I) = \left[1 + \frac{\sum_{i=1}^{511} \min(N_I(i) - N_M(i), 0)}{N_M} \right]$$

If this value exceeds a threshold the object is considered to be present in the image.

The location of the recognised object must now be determined. If $p_{ij}(x,y)$ is the location of pixel j, in bin i of the image histogram, the average location of pixels in bin i will be $L_i(x,y)$. Weighting this by the proportion of pixels of the model histogram M in bin i, and taking the sum of the weighted averages across all the bins of I gives an estimate of the location M(x,y) of the model M in the image I:

$$L_i(x,y) = \frac{\sum_{j=1}^{N_I(i)} p_{ij}(x,y)}{N_I(i)} \qquad M(x,y) = \sum_{i=1}^{511} \left(L_i(x,y) \times \frac{N_M(i)}{N_M} \right)$$

To increase robustness, a background image, showing the workspace with no objects present, is constructed using median filtering. Subtracting the background from each input image means that the histogram containment operator can be restricted to areas of the image most likely to contain an object. While Ballard and Swain (1991) used the familiar red-green-blue colour coding, we build histograms of hue-saturation-intensity values, this representation being more stable under changes in illumination.

To assess the ability of the method to recognise kitchen objects, models were built of six everyday items: a jug, yellow and blue mugs, a sugar pot, coffee jar and a redbush tea packet. Figure 3 shows samples of the image data used to create the model histogram. Each object was placed in the field of view, the background image subtracted and those pixels considered significantly different to the background (i.e. sufficiently likely to depict objects) used to build the histogram. Six recognition trials were then run. In each, one model was loaded into the system and the corresponding object moved across the field of view. Table 1 shows the mean, maximum and standard deviation and maximum values of C(M,I) reported. Values are close to 1 in all cases.

To estimate the level of potential confusion between objects, all six models were then loaded into the system and each object once again presented in turn. Table 2 shows the mean confidence level obtained when each object (vertical axis) is compared with each model (horizontal axis). In most cases maximal values lie along the diagonal. The red tea box is, however, confused with the coffee jar and the jug with the blue mug. Care must be taken when choosing object; the image data reveals similarities not obvious to the naked eye.

Figure 4 shows the system recognising and tracking the positions of two objects simultaneously, the input image is shown in Figure 4a and the system output in figure 4b.

The system has been applied to sequences of images of the stroke patient described above testing the sensor–equipped kitchen objects. Although these early hospital trials were broadly successful, shadows cast by the patient and OT disrupted processing. Large variations in (natural) light also made recognition increasingly difficult as the trial progressed. Future trials will use more controlled illumination. Alternative, illumination independent recognition methods may be considered. Again, development continues.

An early version of the vision system (Ghali et al 2003) was used in isolation, employing a task model expressed in XML to recognise events (sub-tasks) in the hot drink making activity. Although that configuration had some success in determining the relative positions of objects it could not recognise other events, such as an object being tipped. The various embedded sensors can identify (some) binary events but do not record their location. The combination of these complementary technologies is natural. The object recognition/location software is written in C++ over Microsoft DirectShow, and produces a text file listing the name of each object present and its position in image coordinates. These are then converted into the Virtools™ coordinate frame and associated with the corresponding virtual object. Evaluation of the complete tangible user interface is underway.

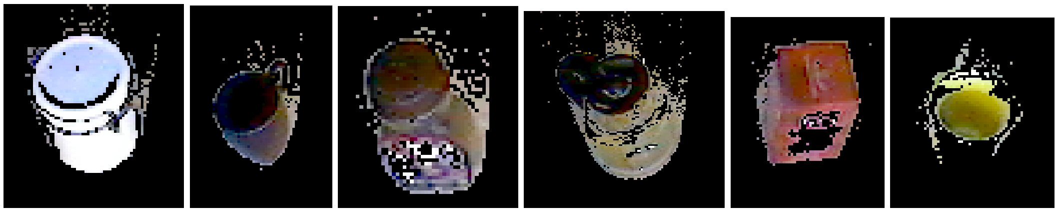


Figure 3. *Background-subtracted image fragments used to build model histograms.*

Table 1. *Confidence scores in a single model/single object trial.*

	JUG	YELLOW MUG	BLUE MUG	SUGAR POT	RED TEA	COFFEE
MAX	0.936198	0.995757	0.994943	0.983957	0.856031	0.979574
MEAN	0.811353	0.925159163	0.888821092	0.830353814	0.6651334	0.81934
STD DEV	0.071397	0.050094091	0.071219554	0.10395138	0.0673606	0.157848

Table 2. *Six object confusion matrix.*

	JUG	YELLOW MUG	BLUE MUG	SUGAR POT	RED TEA	COFFEE
JUG	0.600527955	0.698621582	0.725871254	0.24182894	0.315906881	0.583295731
Y. MUG	0.251440304	0.732824768	0.335289893	0.171731804	0.115175089	0.431838357
B. MUG	0.319896646	0.315210958	0.718631458	0.206522979	0.337775625	0.572546604
SUGAR POT	0.190307208	0.151663981	0.348305208	0.608503226	0.100183472	0.450333453
RED TEA	0.230223298	0.157844021	0.474890362	0.292937234	0.449076957	0.560163957
COFFEE	0.346011333	0.414849035	0.636746737	0.336417579	0.299952175	0.776323632

a. *b.*

Figure 4. *Visual recognition and location of two kitchen objects.*

6. CONCLUSION

The mixed reality systems presented above lie at widely separated points on the continuum of technologies shown in Figure 1, and raise differing technological challenges. In the touch screen system the user/patient interacts directly with virtual objects, focusing attention on the ecological validity of the virtual environment and the extent to which s/he feels immersed in that environment. The tangible interface centres the user's attention on real, physical objects, with the VE providing a task model and a medium for audio-visual feedback rather than a focus for the rehabilitation activity. Manipulation of those objects must, therefore, be smoothly mapped into corresponding changes in the virtual environment. Iterative improvement of these systems, and the design of new systems, will continue over the next period. It should also be recognised that, whilst this combination of technologies may seem cumbersome now, advances in the specialist technologies we are using will improve over the next few years.

We have produced a first version of an integrated mixed reality system that works. Users can access the virtual environment using a variety of interface methods to control activity and complete the task of making a hot drink, with prompts and feedback provided as required. At present, however, this procedure is slow and requires users to learn how to work with it. This is not acceptable for a system intended to support rehabilitation of users who will have very different needs and abilities. Whilst we are aware of the general requirements for stroke rehabilitation, we need to understand more about individual differences amongst the patient population and how this will affect their attitudes towards, and the effectiveness of, the mixed reality system for supporting them through the rehabilitation process. A participatory design approach used for the current phase of development will take information from users (occupational therapists and patients) to directly inform iterative development of new versions of this system.

The longer-term objective is to identify how the various mixtures of the physical and the digital, made available via mixed reality technologies, can and should be exploited during the rehabilitation process:

- Which provide(s) the most appropriate experience(s) at each stage in the rehabilitation process?
- What activities are best supported via these systems?
- Are individual mixed reality systems effective, or is a suite of systems required?

These questions are the focus of clinical evaluation studies conducted by our research team (see Edmans et al 2004) under research currently funded by the Stroke Association.

Acknowledgements. The authors would like to thank their colleagues from the University of Nottingham: Dr Judi Edmans, Dr John Gladman and Dr Marion Walker (Division of Rehabilitation & Ageing), Dr Alan Sunderland and Amanda Porter (School of Psychology) and Dr Danaë Stanton Fraser from the University of Bath.

7. REFERENCES

S.D. Benford, C. Brown, G. Reynard, and C.M. Greenhalgh (1996) Shared spaces: transportation, artificiality and spatiality, *Proc. CSCW '96*, pp 77-85.

M. Billinghurst, S. Weghorst, T. Furness III (1996) Shared space: an augmented reality interface for computer supported cooperative work, *Proc CVE '96*, University of Nottingham, UK.

D.J. Brown, H.R. Neale, S.V.G. Cobband H. Reynolds (1999) Development and evaluation of the virtual city*, International Journal of Virtual Reality, 4*(1), 28-41.

C. Christiansen, B. Abreu, K. Ottenbacher, K. Huffman, B. Massel, and R. Culpepper (1998) Task performance in virtual environments used for cognitive rehabilitation after traumatic brain injury*, Archives of Physical Medicine and Rehabilitation*, 79, 888-892.

S.V. Cobb, T. Starmer, R.C. Cobb, S. Tymms, T.P. Pridmore and D. Webster, D. (2001) Interfacing tactile input devices to a 3D virtual environment for users with special needs

C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, R.V. Kenyon and J.C. Hart (1992) The CAVE: Audio Visual Experience Automatic Virtual Environment, *Communications of the ACM*, 35, 6, pp. 65-72.

R.C. Davies, E. Lofgren, M. Wallergård, A. Linden, K. Boschian, U. Minor, B. Sonesson and G. Johansson (2002) Three applications of virtual reality for brain injury rehabilitation of daily tasks *Proceedings of the 4th International Conference on Disability Virtual Reality and Associated Technologies* Veszprem, Hungary.

J.A. Edmans, J. Gladman, M. Walker, A. Sunderland, A. Porter. and D. Stanton Fraser (2004) Mixed reality environments in stroke rehabilitation: development as rehabilitation tools, *Proceedings of the 5th International Conference on Disability Virtual Reality and Associated Technologies* Cambridge,UK.

A. Ghali, A. Cunningham and T.P. Pridmore (2003) Object and event recognition for stroke rehabilitation, *Proc. Visual Communication and Image Processing*, Lugano, Switzerland.

D. Gourlay, K.C Lun, Y.N. Lee and J. Tay (2000) *Virtual Reality*.

D. Hilton, S.V. Cobb, T.P. Pridmore (2000) Virtual reality and stroke assessment: therapists perspectives *Proceedings of the 3rd International Conference on Disability Virtual Reality and Associated Technologies*, Alghero, Sardinia.

D. Hilton, S.V. Cobb, T.P. Pridmore and J. Gladman (2002) Virtual reality and stroke assessment: a tangible interface to an every day task *Proceedings of the 4th International Conference on Disability Virtual Reality and Associated Technologies* Veszprem, Hungary.

M. Hofmann, A. Rosier W. Schwarz F. Muller-Spahn, K. Krauchi, C. Hock and E. Seifritz (2003) Interactive computer training as a therapeutic tool in Alzheimer's disease, *Comprehensive Psychiatry*, 44, 3, pp 213-219.

H. Ishii and B. Ullmer (1997) Tangible bits: towards seamless interfaces between people, bits and atoms, *Proc. CHI '97*, pp 234-241..

H. Ishii and B. Ullmer (2001) Emerging frameworks for tangible user interfaces. In *Human computer Interaction in the new millenium*, Carroll J (Ed) Addison Wesley, pp 579-601.

B.N. Koleva S.D. Benford, and C.M. Greenhalgh (1999) The properties of mixed reality boundaries, *Proc. 6th European Conference on Computer Supported Co-operative Work,* Copenhagen, pp. 119-137

A. Linden, R.C. Davies, K. Boschian, U. Minör, R. Olsson, B. Sonesson, M. Wallergard, G. Johansson (2000) Special considerations for navigation and interaction in virtual environments for people with brain injury. *Proceedings of the 3rd International Conference on Disability: Virtual Reality and Associated Technologies* Alghero, Sardinia.

Mendozzi et al. (2000) VIRT –factory trainer project: a generic productive process to train persons with disabilities. *Proc 3rd Intl Conf on Disability: Virtual Reality and Assoc Technologies*, Alghero, Sardinia.

P. Milgram and F. Kishino (1994) A taxonomy of mixed reality visual displays, *IEICE Trans. Information Systems*,, Vol E77-D, 12.

R. Potter, J. Weldon, and B. Schneiderman (1988) Improving the accuracy of touch screens: an experimental evaluation of three strategies, *Proc. CHI '88*, pp 27-32.

J.W. Regian, W.L. Shebilske and J.M. Monk (1992) Virtual Reality: An Instructional Medium for Visuo-Spatial Tasks. *J Comm* 4, 136-149.

Rose, Brooks & Attree (2000) Virtual reality in vocational training of people with learning disabilities, *Proc 3rd Intl Conf on Disability: Virtual Reality and Assoc Technologies*, Alghero, Sardinia.

E. Sharlin, Y. Itoh, Y., Watson, B., Kitamura, Y., Sutphen, S., and L. Liu (2002) Cognitive cubes: A tangible use interface for Cognitive assessment, *Proc. CHI '2002* pp 347-354.

D. Stanton et al. (2000) Virtual environment as spatial training aids for children and adults with physical disabilities. *Proc 3rd Intl Conf on Disability: Virtual Reality and Assoc Technologies,* Alghero, Sardinia.

M.J. Swain and D. H. Ballard (1991) Color indexing, *International Journal of Computer Vision*, 7, 1, pp. 11-32.

P.L. Weiss, Y. Naveh, and N. Katz (2003) Design and testing of a virtual environment to train stroke patients with unilateral spatial neglect to cross a street safely. *Occupational Therapy International* 10(1) pp 39-55.

Virtual reality based intervention in rehabilitation: relationship between motor and cognitive abilities and performance within virtual environments for patients with stroke

R Kizony[1], N Katz[2] and P L Weiss[3]

[1,2]School of Occupational Therapy, Hadassah-Hebrew University, Jerusalem, ISRAEL

[1,3]Department of Occupational Therapy, University of Haifa, Haifa, ISRAEL

[1]*rachelk@zahav.net.il*, [2]*noomi.katz@huji.ac.il*, [3]*tamar@research.haifa.ac.il*

ABSTRACT

The objective of this study was to provide experimental data to support a proposed model of VR-based intervention. More specifically our goal was to examine the relationships between cognitive and motor ability and performance within virtual environments. Thirteen participants who have had a stroke participated in the study. They each experienced three virtual environments (Birds & Balls, Soccer and Snowboard) delivered by the GX- video capture system. After each environment they complete a scenario specific questionnaire and Borg's scale for perceived exertion. Their cognitive, motor and sensory abilities were measured as well. The participants' responses to the VR environments showed that they enjoyed the experience and felt high levels of presence. The results also revealed some moderate relationships between several cognitive abilities and VR performance. In contrast, the motor abilities and VR performance were inversely correlated. In addition, there was a relationship between presence and performance within the Soccer environment. Although these results support some components of the proposed model it appears that the dynamic nature of the virtual experiences would be more suited to comparisons with different measures of motor ability than those used in the current study.

1. INTRODUCTION

Stroke is a major cause of disability for adults and the elderly, often resulting in motor and cognitive impairment, and functional disability (Woodson, 1995). Various studies have found relationships between functional disability and both cognitive and motor deficits. Thus a major goal of the rehabilitation process is to improve these deficits (Katz, et al.1999).

Virtual Reality (VR) has recently begun to be used for rehabilitation of patients with stroke. Piron, et al. (2001) used VR to train reaching movements, Broeren, et al., (2002) developed a VR haptic device for the assessment and training of motor coordination and Jack et al. (2001) and Merians et al. (2002) used a force feedback glove to improve range of motion, speed and strength of hand movement. The results of the latter study, which included three patients with stroke, showed VR to be useful for the improvement of upper extremity function in patients who are at a chronic stage. VR also appears to be beneficial for the training of safe street crossing using a desktop platform with patients who suffer from unilateral spatial neglect (Katz et al., 2003; Weiss et al., 2003).

The virtual reality experience is multidimensional and appears to be influenced by many parameters whose interactions remain to be clarified. A proposed model for virtual reality in rehabilitation is presented in Fig. 1. This model was developed within the context of the International Classification of Functioning, Disability and Health (ICF) (World Health Organization, 2001) terminology (Weiss et al., In Press) and consists of three nested circles, the inner "Interaction Space", the intermediate "Transfer Phase" and the outer "Real World".

As represented schematically in Fig. 1, two primary factors within the "Interaction Space" influence the nature of the interaction between the user and the virtual environment. The first of these factors relates to the user's personal characteristics. These include demographic factors (e.g. age, gender, cultural background), body functions (e.g. cognitive, sensory, motor) and structures (e.g., arms, legs). The second factor relates to characteristics of the virtual environment including both the type of VR platform and its underlying technology (enabling the flow of information to and from the user) and the nature and demands of the task to be performed within the virtual environment. The characteristics of the virtual environment may be either barriers or enablers to performance. The client interacts within the virtual environment, performing functional or game-like tasks of varying levels of difficulty. This enables the therapist to determine the optimal environmental factors for the client. Within the "Interaction Space" sensations and perceptions related to the virtual experience take place; here the user's sense of presence is established, and the process of assigning meaning to the virtual experience and the actual performance of virtual tasks or activities occur.

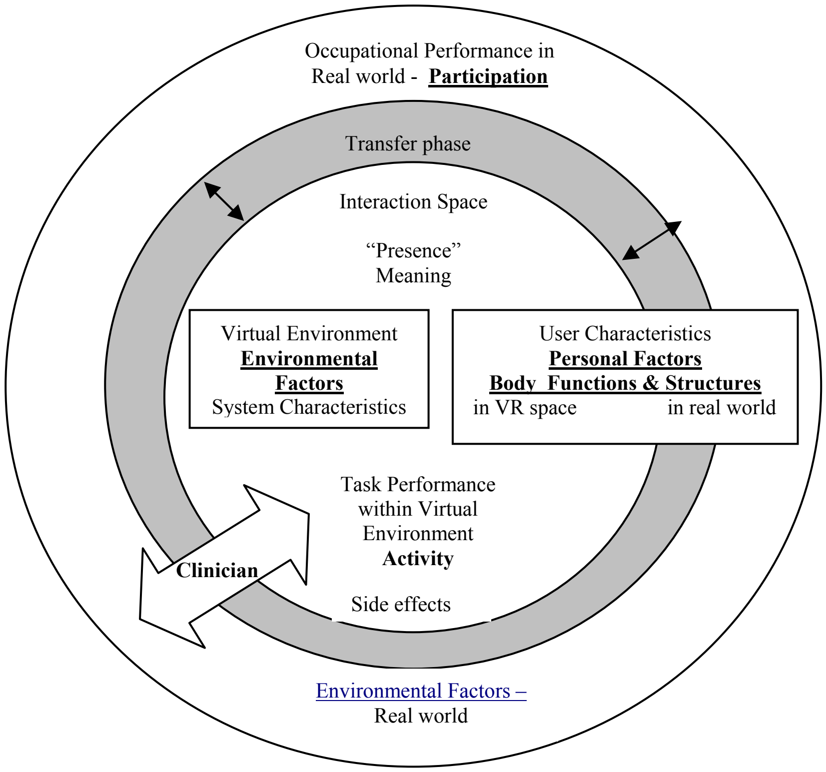


Figure 1: *A model of VR-based rehabilitation within the context of terminology from the International Classification of Functioning, Disability and Health concepts (indicated in bold).*

One of the undesired consequences of interacting within some virtual environments that could affect performance are side effects such as nausea and dizziness (Kennedy, & Stanney, 1996). For those users who are susceptible to this problem, such effects usually occur only in certain types of VR systems such as when using an HMD (Rand, et al., in press). Clinicians should be aware of this possible effect and ensure that clients who are susceptible to side effects avoid using a VR system and environment for prolonged periods of time.

From the Interaction Space (inner circle) we move to the Transfer Phase (intermediate circle) since our goal in rehabilitation is to improve daily function in the real world and this requires transfer of the trained skills or tasks as well as environmental modifications from the virtual environment to the real world. The "Transfer Phase" may be very rapid and accomplished entirely by the client or may take time and need considerable guidance and mediation from the clinician. Finally, the large, outer circle represents real world environments illustrating that the ultimate goal is to help the client achieve participation in the real world environment by overcoming, adapting to or minimizing the environmental barriers. The entire process is facilitated by the clinician whose expertise helps to actualize the potential of VR as a rehabilitation tool.

A key principle in rehabilitation is matching therapeutic tasks to the patients' abilities in order to enable them to improve residual capabilities without causing fatigue and frustration. Knowledge about the relationship between user characteristics, sense of presence and performance within the virtual environment would help clinicians achieve an optimal match, enabling them to select and customize environments and tasks so that they are more suited to patients' abilities. This would also enhance patients' involvement in the task. The application of this principle to people with stroke is particularly important since their disability is complex with both motor and cognitive components. Thus, a better understanding of the relationship between deficit and performance within virtual environments should lead to greater efficiency in the use of VR-based intervention.

The objective of this study was to provide experimental data to support the proposed model of VR-based intervention. We begin by an examination of the relationships between motor, sensory and cognitive abilities to performance within three virtual scenarios which, as indicated above, is an essential step for understanding the nature of interaction within the virtual environment.

2. METHODS

2.1 Participants

In order to establish the protocol and inclusion criteria for the current study, we first tested three participants with a protocol similar to one used in a prior study with patients who had spinal cord injuries (Kizony et al., 2003). Thirteen patients with stroke (4 female, 9 male) with a mean age of 66.3 ± 7.9 participated in the full study. Seven participants had a left hemispheric stroke and six had a right hemispheric stroke. Time between onset of the stroke and participation in the study ranged from 5 weeks to 11 months. All participants took part in rehabilitation services, either in hospital (11) or at an ambulatory day centre (2). Inclusion criteria included independence in ADL prior to the stroke, the ability to understand instructions and sign the informed consent and to move the affected upper extremity independently or with the aid of the non-affected arm.

2.2 Instruments

2.2.1 Virtual Reality

1. VividGroup's GX video-capture VR System has potentially important applications for the rehabilitation of children and adults with physical and/or cognitive impairment [www.vividgroup.com, www.irexonline.com] (Kizony, et al., 2003a,b; Cunningham & Krishack, 1999; Sveistrup, et al., 2003). The advantages of using this VR system with patients with brain damage has been described in detail elsewhere (Kizony et al., 2003a,b; Kizony et al., 2002) and includes its natural way of interaction and very low incidence of sides effect. Three virtual environments were used with this system:

Birds & Balls: The user sees himself in a pastoral setting and touches virtual balls (Fig. 2a). Performance in this environment was measured in terms of percent success (i.e., the number of balls touched out of the total numbers of balls) and response time (i.e., the time between the ball's first appearance on the screen until it was touched).

Soccer: The user is a goalkeeper and has to prevent the balls from entering the net (Fig. 2b). The performance in this environment was measured in terms of percent success similar to Birds & Balls.

Snowboard: The user is required to avoid obstacles as he skis downhill. Performance here was measured in terms on percent success (i.e. the number of obstacles avoided out of the total number of obstacles).

Figure 2a. *A participant within the Birds & Balls environment*

Figure 2b. *A participant within the Soccer environment*

For Birds & Balls and Snowboard, the third minute of each game was analyzed since it should reflect the participant's best performance, i.e., after participants had practiced but prior to the onset of fatigue. For Soccer, the second minute was analyzed since not all the participants were able to continue to perform at the same level during the third minute due to their motor impairments.

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

2. The Scenario Feedback Questionnaire (SFQ) is a 8-item questionnaire designed to obtain information about the subjective responses of the participants to the VR experience in each scenario. It queries the user's sense of presence, perceived difficulty of the task and any discomfort that users may have felt during the experience. The first six items of the questionnaire were formulated as an abbreviated alternative to the longer Presence Questionnaire developed by Witmer and Singer's [20]. These items assessed the participant's (1) feeling of enjoyment, (2) sense of being in the environment, (3) success, (4) control, (5) perception of the environment as being realistic and (6) whether the feedback from the computer was understandable. The seventh item, queried whether participants felt any discomfort during the experience. A eighth item queried their perceived difficulty while performing the task. Responses to the first seven items were rated on a scale of 1-5 where 1=not at all and 5 = very much. Responses to the eighth item was also rated on a 1-5 scale where 1= very easy and 5 = very difficult.

2.2.2 Motor and Sensory Abilities. Active movement and coordination of the affected upper extremity was tested using the Fugl-Meyer Motor Assessment (Fugl-Meyer et al., 1975), muscle tone was tested with the modified Ashworth scale (Bohannon & Smith, 1987), touch sensation was tested and proprioception was tested with the Thumb test (Prescott et al., 1982). Balance was measured via the Functional Reach Test (Duncan et al., 1990) with its modification for sitting (Lynch et al., 1998). For this test we calculated a total score which is the sum of leaning forward when the left side is near the wall and leaning forward when the right side is near the wall. The perceived exertion during each task was measured via Borg's (1990) scale (scores range between 6 (no exertion) and 20 (maximal exertion)).

Cognitive abilities: Tests for visual search and attention included the Star cancellation test (Wilson et al., 1987) and the Mesulam symbol cancellation test (Weintraub et al., 1987). For each of these tests we calculated the number of correct symbols cancelled as well as the time it took to complete the test. Visual memory was tested with the Contextual Memory Test (Toglia, 1993), and thinking operations (i.e. categorization and sequence) with the Lowenstein Occupational Therapy Cognitive Assessment (Itzkovich et al., 2000) and executive functions with the Behavioral Assessment of Dysexecutive Syndrome (Wilson et al., 1996). (The latter results are not reported in the present paper.)

2.3 Procedure

Participants experienced each of the three virtual environments for 3-4 minutes, depending on fatigue. After each environment the participant completed the SFQ and the perceived exertion scale. The participant's motor and cognitive abilities were evaluated within approximately one week of the VR session.

2.4 Data Analysis

Descriptive statistics were used to describe the participants' responses to the SFQ, their performance within the virtual environments and their cognitive and motor scores. To examine differences between the scenarios regarding percent success, perceived exertion and difficulty we used either paired t-tests or Wilcoxon tests for related groups respectively. To examine relationships between virtual performance and cognitive abilities, motor abilities and sense of presence, Pearson or Spearman correlations were used depending on the nature of the scale of measurement (i.e. ordinal or interval). Due to the small sample, marginally significant values (e.g., $p<.053$) are presented as well.

3. RESULTS

3.1 Feedback on the VR Experience

The participants expressed their interest in having this therapy; indeed, the majority requested and received additional VR sessions following their completion of the study. Their responses to the SFQ showed that they enjoyed the experience and felt high levels of presence for the different environments with a mean ± standard deviation for Birds & Balls equal to 4.4 ± 0.4, Soccer equal to 4.1 ± 0.7, and Snowboard equal to 4.3 ± 0.5 out of a total score of 5. In addition, no cybersickness-like side effects were reported.

3.2. Performance within the Virtual Environments and Sensorimotor and Cognitive Tests

The participants' performance within the three environments is presented in Table 1. The participants performed significantly better in Snowboard than in Birds & Balls ($t=-3.5;p=.004$) and in Soccer ($t=-6.7$; $p=.000$). Performance in Birds & Balls was significantly better than in Soccer ($t=4.2;p=.001$). Their perceived exertion was significantly higher for Soccer then for Birds & Balls (t=-.2.43; p=032). The other differences for exertion were not significant. Perceived difficulty (from the SFQ) was significantly higher in Soccer than in Snowboard (z=-2.25; p=.024). The other differences for difficulty were not significant.

Table 1. *Performance within the three virtual environments (N = 13)*

	Birds & Balls Level 2[a]	Soccer Level 1[b]	Snowboard
Percent success	75.7 ±17.0	53.7 ± 22.8	91.1 ±6.2
Response time	5.7 ± 1.4	N/A	N/A
Exertion	9.3 ± 2.3	10.9 ± 2.7	9.5 ± 2.6
Perceived difficulty	2.0 ± 1.0	2.6 ± 1.1	1.8 ± 0.9

a. For Birds & Balls, level 2, four balls simultaneously approach the participant from different directions

b. For Soccer, level 1, one to two balls approach net simultaneously from different directions

Table 2 shows mean performance scores on the motor and cognitive tests. Participant scores on the cognitive tests were high with a relatively small variance. Although the variance was larger on the motor measures, ten out of the 13 participants had functional active movements in their affected upper extremity and sufficient balance to be able to reach out from their midline. Nine participants had increased muscle tone and four had normal muscle tone as measured during elbow flexion and extension movements. In the sensory measures it was found that 11 participants had intact touch sensation but only eight had intact proprioception.

Table *2*. *Performance on the cognitive and motor measures (N = 13)*

Measure	Mean ± SD	Range of possible scores
Cognitive		
Star Cancellation score	52.4 ± 2.1	0 - 54
Star Cancellation time (s)	107.8 ± 98.2	
Mesulam score	56.1 ± 5.0	0 - 60
Mesulam time	167.2 ± 86.4	
Contextual Memory Test	19.7 ± 7.4	0 - 40
Categorization	3.3 ± 0.8	1 - 4
Sequence	3.1 ±1.1	1 - 4
Motor and Sensory		
Functional Reach Test	70.0 ± 19.1	
Fugl- Meyer active movements	44.7 ± 15.5	0 - 60
coordination / speed	3.8 ± 1.2	0 - 6

3.3 Relationships between Sense of Presence and Enjoyment and Performance within Virtual Environment

Significant correlations were found between percent success in Soccer and the sense of presence as measured by the SFQ (r=.56; p=.05) and level of enjoyment also measured by the SFQ (r=.59; p=.03). The remaining correlations between presence, enjoyment and performance were not significant.

3.4 Relationships between Cognitive and Motor Abilities and Performance within the Virtual Environments

Significant correlations were found between several of the cognitive tests and performance within the virtual environments. The categorization test was correlated with both response time in Birds & Balls (r=-.61; p <.05) and percent success in Soccer (r=.57; p=.053). The Mesulam attention score was correlated with percent success in Snowboard (r=.56; p<.05) and the CMT memory test with percent success in Birds & Balls (r=.57; p<.05). These correlations indicate that the greater the cognitive ability, the better the performance in the virtual environments. With regard to the motor and sensory tests, the only significant correlations were found between the Fugl-Meyer coordination/speed test and percent success in Soccer (r=-.84; p=.001) and response time in Birds & Balls (r= 59; p=.54). Surprisingly, these correlations indicate that better motor ability is related to poorer performance in the virtual environments.

4. DISCUSSION

The model presented in the Introduction describes an Interaction Space which refers to a participant's performance within a virtual environment. Given appropriate conditions (e.g., feedback of sufficient quantity and quality), participants will feel "present" within the virtual world and it will become meaningful to them. The responses of participants in the current study attest to attainment of these two important reactions. The results of the short presence scores were similar to those of the patients with spinal cord injuries tested previously (Kizony et al., 2003b); their involvement within the environment appeared to encourage them to participate in what would otherwise be painful or boring therapeutic activities. There is already considerable evidence demonstrating that many VR-based interventions motivate routine therapy (Jack et al, 2001; Sveistrup, et al., 2003). The presence and enjoyment scores recorded in the present study provide additional evidence of this important VR asset (Rizzo et al., 2004). It also appears that, in addition to their motivational characteristics, these games encouraged participation by providing activities that were meaningful to the participants. . One participant remarked: "I like soccer the best since it reminds me of my grandson who plays soccer in a professional youth team".

Performance in an activity within the Interaction space is carried out within the context of a meeting between participants' personal characteristics and those of the task and environment. Thus one of the goals of the current study was to examine the relationships between cognitive and motor abilities and performance within the virtual environment. Although these results must be interpreted with caution due to the small sample size and the small variance of the tests scores, nevertheless, there appears to be a relationship between categorization which requires abstract thinking, visual contextual memory and visual attention and improved performance in the virtual environment. This last finding is in accordance with the nature of the virtual tasks used in this study which require participants to rely heavily on visual attention (i.e., searching for and responding to stimuli coming from all directions). The proposed model highlights the impact that user characteristics, such as ability to engage in visual attention tasks, have on performance within the virtual environment. It would therefore appear important to screen for such abilities prior to selecting a given virtual scenario. In order to further test the importance of user characteristics, we are currently analyzing the relationships between higher cognitive functions such as executive functions and VR performance.

In contrast to cognitive abilities, there were very few correlations between motor abilities and performance within the virtual environments. This was perhaps due to the small variance in this sample with most of the participants having higher level motor abilities (i.e., an insufficient data spread to obtain adequate correlations). One of the motor ability/virtual performance correlations appeared to be anomalous. That is, participant coordination/speed was found to be inversely related to performance within the virtual environments. This may have been due to the way in which motor ability was tested relative to the way it was performed virtually; the test task required that the participant touch his nose with the affected arm in a controlled and precise manner. In contrast, movement within the virtual scenarios used in this study, especially in Soccer due to the speed of the approaching balls, entailed ballistic actions. Second, touching the targeted balls did not require precise movements. Moreover, participants were required to lean forward and to the side within the virtual scenario, movement types which, again, differed greatly from the tests of motor and sensory abilities. Finally, the motor/coordination test rates ability only of the affected upper extremity. In contrast, within the virtual scenarios, the participants occasionally used their unaffected arm as well as other body parts to interact. In contrast to the pre-VR tests of functional ability, the virtual experience is clearly dynamic, and entails the use of many cognitive and motor abilities simultaneously. In retrospect, it is clear that additional measures of motor ability are required in order to more accurately characterize the relationship between it and performance within the virtual environment.

The above discussion leads to an additional important question - should we expect performance demands, and hence their characteristics, within the real and the virtual worlds to be identical? It may be that differences in presence, motivation, or other factors influence the movement patterns. Viau et al (2004) found that the movements in a virtual task were similar to those in a comparable task performed in the real world. In contrast, Lott et al. (2003) found there to be significant differences between functional lateral reach when performed in the real environment versus a virtual environment delivered with the GX video capture system; in this case, movements in the virtual environment were of higher quality. We recommend that, in future, additional data on cognitive, motor as well as functional abilities be measured on a larger number of participants in order to perform multiple regression which will help to predict and explain the virtual performance.

Further data will also help to clarify the relationship between presence and performance which takes place within the Interaction space. The results of the present study provided some support for a relationship between presence and performance via the results from the Soccer game. This environment is the most difficult one to perform in, a fact that perhaps compels a participant to "be there" in order to succeed.

In addition to the Interaction space, the model also designates a Transfer Phase and a Real World space. Since the ultimate goal of therapy is to enable individuals to transfer the skills learned during rehabilitation to adaptive performance in the real world, it is essential that these two components be explicitly tested. Initial positive results showing the possible transfer of skills to the real world have been described by Jack et al. (2001) as cited above. We are currently carrying out such a study, using a single subject design on patients with stroke.

5. REFERENCES

R W Bohannon and M B Smith (1987), Interrater reliability of a modified Ashworth scale of muscle spasticity. Phy. Ther., 67, pp. 206-207.

G Borg (1990), Psychophysical scaling with applications in physical work and the perception of exertion, *The Scand. J. Work Environ. Health, 16 Suppl 1*, pp. 55-58.

J Broeren, A Bjorkdahl, R Pascher and M Rydmark (2002), Virtual reality and haptics as an assessment devise in the postacute phase after stroke. *CyberPsych. Behav., 5,* pp. 207-211.

D Cunnigaham and M Krishack (1999), Virtual reality promotes visual and cognitive function in rehabilitation. *CyberPsychol. and Behav., 2,* pp. 19-23.

P W Duncan, D K Weiner, J Chandler and S Studenski (1990), Functional reach: a new clinical measure of balance, *J. Gerontol. 45,* pp. M192-197.

A R Fugl-Meyer, L Jaasko, I Leyman, S Olsson and S Steglind (1975), The post stroke hemiplegic patient: a method of evaluation of physical performance, *Scand. J. Rehab. Med., 7,* pp. 13-31.

M Itzkovich, B Elazar, S Averbuch and N Katz (2000), *LOTCA Manual*. (2nd edition). New Jersey: Maddak.

D Jack, R Boian, A Merians, M Tremaine, G Burdea, A Adamovich, M Recce and H Poizner (2001), Virtual Reality-Enhanced Stroke Rehabilitation, IEEE Trans. *Neural Sys. Rehabil. Eng., 9,* pp. 308-318.

N Katz, A Hartman-Maeir, H Ring and N Soroker (1999), Functional disability and rehabilitation outcome in right hemisphere damaged patients with and without unilateral spatial neglect. *Arch. Phy. Med. Rehab. 80,* pp. 379- 384.

N Katz, H Ring, Y Naveh, R Kizony, U Feintuch and P L Weiss (2003), Effect of interactive virtual environment training on independent safe street crossing of stroke patients with unilateral spatial neglect. *Abstract for the Israeli Association of Physical Medicine and Rehabilitation*. Shfayim, Israel.

R S Kennedy and K M Stanney (1996), Postural instability induced by virtual reality exposure: Development of certification protocol. *Intl. J. Human-Computer Inter., 8*, pp.2547.

R Kizony, N Katz, H Weingarden and P L Weiss (2002), Immersion without encumbrance: adapting a virtual reality system for the rehabilitation of individuals with stroke and spinal cord injury, *Proc. 4th Intl Con. Disabil. Virtual Reality Assoc. Technol.* University of Reading: Vresprem, Hungary, pp 55-61.

R Kizony, N Katz and P L Weiss (2003a). Adapting an immersive virtual reality system for rehabilitation. *J.Visual. Comp. Anim. 14*, pp. 261-268.

R Kizony, L Raz, N Katz, H Weingarden and P L Weiss (2003b), Using a video projected VR system for patients with spinal cord injury. *Proc. 2nd Intl. Workshop Virtual Rehab.* Rutgers University: New Jersey, USA, pp 82-88.

A Lott, E Bisson, Y Lajoie, J McComas and H Sveistrup (2003), The effect of two types of virtual reality on voluntary center of pressure displacement. *Cyberpsychol. Behav.*, *6,* pp. 477-485.

S M Lynch, P Leahy and S P Barker (1998), Reliability of measurements obtained with a modified functional reach test in subjects with spinal cord injury, *Phys. Ther. 78*, pp. 128-133.

A Merians, D Jack, R Boian, M Tremaine, G C Burdea, S V Adamovich, M Recce and H Poizner (2002), Virtual reality- augmented rehabilitation for patients following stroke. *Phys. Ther. 82,* pp. 898-915.

L Piron, F Cenni, P Tonin and M Dam (2001), Virtual Reality as an assessment tool for arm motor deficits after brain lesions. *Stud. Health Techno. Inform. 81*, pp. 386-392.

R J Prescott, W M Garraway and A L Akhtar (1982), Predicting functional outcome following acute stroke using a standard clinical examination, *Stroke, 13,* pp. 641-647.

D Rand, R Kizony, U Feintuch, N Katz, N Josman, AA Rizzo and P L Weiss (In press), Comparison of two VR platforms for rehabilitation: Video capture versus HMD, *Presence*.

A A Rizzo, M T Schultheis, K Kerns and C Mateer (2004), Analysis of assets for virtual reality in neuropsychology, *Neuropsychol. Rehab. 14,* pp. 207-239.

H Sveistrup, J McComas, M Thornton, S Marshall, H Finestone, A McCormick, K Babulic and A Mayhew (2003), Experimental Studies of Virtual Reality-Delivered Compared to Conventional Exercise Programs for Rehabilitation. *CyberPsychol. Behav., 6,* pp. 245 – 249.

J P Toglia (1993), *Contextual Memory Test manual*. Arizona:Therapy Skill Builders.

A Viau, M F Levin, B J McFadyen and A G Feldman (2004), Reaching in reality and in virtual reality: a comparison of movement kinematics. *Proc. 15th Cong. of the Intl. Soc. Of Electrophysiology and Kinesiology,* Bostom University, USA. pp. 50.

S Weintraub and M M Mesulam (1987), Right cerebral dominance in spatial attention: further evidence based on ipsilateral neglect. *Arch. Neurol. 44,* pp. 621-625.

P L Weiss, R Kizony, U Feintuch and N Katz (In Press), Virtual reality in neurorehabilitation. In *Textbook of Neural Repair and Neurorehabilitation*. (M E Selzer, L Cohen, FH Gage, S Clarke & P W Duncan Eds.). Cambridge Press, In Press.

P L Weiss, Y Naveh and N Katz (2003), Design and testing of a virtual environment to train CVA patients with unilateral spatial neglect to cross a street safely. *Occup. Ther. Intl. 10*, pp.39-55.

B A Wilson, N Alderman, P W Burgess, H Emslie and J J Evans (1996), *Behavioral Assessment of the Dysexecutive syndrome.* Bury St. Edmunds: Thames Valley Test Company.

B A Wilson, J Cockburn and P W Halligan (1987), *The Behavioral Inattention Test*. England: Thames Valley Test Company.

A M Woodson (1995), Stroke. In *Occupational therapy for physical dysfunction. (*C A Trombly Ed.) (4th ed.). (pp.677-704). Baltimore: Williams & Wilkins.

World Health Organization. (2001). International *classification of functioning disability and health (ICF).* Geneva: World Health Organization.

Evaluation of a computer-assisted 2D interactive virtual reality system in training street survival skills of people with stroke

Y S Lam[1], S F Tam[2], D W K Man[3] and P L Weiss[4]

[1,2,3]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, HONG KONG

[4]Faculty of Social Welfare & Health Studies, University of Haifa, ISRAEL

[1]02901381r@polyu.edu.hk, [2]rsalan@polyu.edu.hk, [3]rsdavid@polyu.edu.hk, [4]tamar@research.haifa.ac.il

ABSTRACT

The rationale, procedures and results of pilot study of the VR system in training street survival skills of people with stroke were presented and discussed. The following main study was also refined from the outcomes to make it more feasible and potentially beneficial to the patients.

1. INTRODUCTION

Stroke is a focal or global neurological impairment of sudden onset, and lasting more than 24 hours (or leading to death) and of presumed vascular etiology (WHO, 1978). Resulted cognitive deficiencies represent substantial sources of disability. The functional limitations in the activities of daily living, work and leisure could be resulted from those impaired functioning. Therefore, a systematic functionally oriented services of cognitive rehabilitation based on assessment and understanding of individual deficits were used to compensate for the cognitive deficits and promote functional adaptation in daily living (Gontkovsky et al, 2002). The traditional methods have involved the various paper-and-pencil tasks, manipulation of blocks and real-life activities. Learning theory, cognitive theory and neuropsychology are used to direct the development of the treatment and explain the results. Since the technological developments over the past few years had led to the use of computers in the rehabilitation, there are numerous computer-based cognitive rehabilitation programs for different purpose. Particularly, using Virtual Reality (VR) as a medium in rehabilitation has becoming more popular and was progressing in a rapid way.

Virtual Reality (VR) can be viewed as an advanced computer interface that allows the user to interact and become immersed within computer-generated simulated environments (Rizzo et al. 1997). Rapid development in modern technology and sophisticated computer systems have made it possible for desktop computers to display complex visual images that change in response to instructions from users (Christiansen et al. 1998). According to Rose et al. (2000), the virtual environments demonstrate many of the characteristics of an ideal training medium. The virtual environments are especially valuable when training in real life situations will be impractical, dangerous, logistically difficult, unduly expensive or too difficult to control. The person can actively interact with the simulated world by using interface devices. Its realism and versatility makes VR a suitable and innovative approach in rehabilitation. A number of researches have supported the use of VR in rehabilitation, both assessments and treatments, among physical, cognitive and psychological conditions. The immerse VR can offer higher level of presence and realism while the non-immerse one is more affordable and accessible to patients due to simpler computer requirements and more suitable for the patients feeling nausea in the immerse VR. The essence of VR gives the person a sense of immersion and presence in a virtual environment but like acting in real world (Hoffman et al, 2001). Since VR is a kind of advance computer applications, it has all the advantages in other software programs such as modifiability, adaptability, customizability and portability (Rizzo et al, 1997). Therefore, with the appropriate programming, VR can create a world with specific purpose. Currently, VR has been used successfully to train flight pilot, surgeons and treat patients with phobias while the use of VR in rehabilitation is comparatively limited (Campbell, 2002; Christiansen, et al, 1998; Grealy et al, 1999; Olasav et al, 2002; Rizzo et al, 1997).

2. MAIN STUDY

The present study will evaluate the effectiveness of a newly developed 2D non-immerse VR in training street survival skills of people with cognitive deficits. Besides street survival skills training, the present 2D VR strategy can also be used in assessing and training various community survival skills like transport skills, road safety skills, wheelchair accessibility, etc. Such a rehabilitation development would complement hospital-based rehabilitation and community services, and might guarantee success in achieving a better quality of life. A 2DVR based street survival skills training program and a corresponding psycho-educational package will be developed for the present study. The training contents are the concepts of street survival skills including road crossing, use of MTR and money management.

2.1 Research Objectives

The present study will develop and compare the effectiveness among a non-immerse flat screen 2DVR method, a conventional psychoeducational method and control group in training street survival skills of people with stroke. The project objectives are:

- To develop a 2DVR system as a training modality for street survival skills
- To evaluate the effectiveness of a 2DVR system as a training modality for street survival skills in terms of behaviours, self-efficacy and functional performance
- To compare the effectiveness of the virtual reality-based and conventional psycho-educational strategies in training street survival skills

2.2 Design

A randomized control group pre-test and post-test design will be adopted. The subjects will be randomly assigned into 3 treatment groups. 20 subjects will receive street survival skills training using a newly developed 2D VR based computer program and learn the skills through interaction within an interactive virtual environment. Another 20 subjects will receive the same skills training using a psycho-educational package using demonstration, role-play, and immediate feedback with verbal reinforcement. The remained 20 subjects will be in the control group. The measurement will be taken at the beginning of the training and the end of the training. There are 10 sessions and each last 1-hour. Each treatment group will be subdivided into 4 smaller groups of 5 subjects for the ease of training in terms of the equipment, the space and effective group size. The pilot study was done prior to the main study in which the treatment contents, the instructional strategies and the outcome measures can be tested and refined to formulate a better research design with lesser flaws and more effective treatment results. The usability testing and the qualitative evaluation will also be performed in addition to the quantitative measurement. These useful information and valuable feedback from the subjects can be used to modify the research design in the main study, so that the confounding or extraneous variables can be minimized to improve the reliability and the validity of the research study.

2.3 Sampling

The blocking is adopted in the sampling process. In the stage 1, 4 local organizations are selected and contacted for the potential population of stroke patients. In the stage 2, the patients within each block are randomly allocated into 3 groups by random assignment. The randomization is done by using the random number generation in SPSS 10. Each subject has an equal chance of being assigned to any group, impendent of personal judgment and bias. A total of 60 subjects with strokes will be recruited. At a sample size of 60, an alpha of 0.05, the number of level (k) of 3 and an expected effect size of 0.5 (medium effect size) for the interventions, the power is 0.93 (Cohen, 1987) in the statistical power analysis for the behavioral sciences.

2.4 Selection criteria

- Persons age 40 to 70 who suffer from stroke
- Medically and emotionally stable
- Able to follow simple instruction, write with a pen in Chinese or English
- Have good visual tracking, discrimination and figure-ground skills
- Have sustained attention span at least 10 minutes
- No previous psychiatric history and no computer-phobia

2.5 Outcome Measures

A behavioral checklist will be developed according to the results of content-specific task analysis. The performances of the skills are broken down into 22 small steps that can be rated from 0 (dependent) to 3 (independent). The subjects will go through the same assessment procedures in the real setting before, after and at one-month follow-up. The checklist is used to obtain the baseline performance, the initial performance and the retention performance. Therefore, it can be made valid and reliable comparisons among the groups and within the group as well as the measurement of the retention of the skills and the generalization of the skills into the real life situation.

Self-efficacy scale will be developed according to Bandura's principles (1997) for the study. It is a measurement of self-perception in the ability of specific tasks. The scale was created to assess a specific sense of perceived self-efficacy with the aim in mind to predict coping with street survival as well as adaptation after experiences in different kinds of life events. The scale is self-administered and comprehensive questionnaire. The patients mark their perceived self-efficacy on a 10-point Likert scale. The scoring is done by summing up the responses to all 11 items to yield the final composite score with a range from 10 to 110. The higher scores mean better self-efficacy.

Formative evaluation will be done throughout the study by observations in behaviours, verbal expressions, comments, feelings, participation during the treatment. Both close-end and open-end questions will be used in the qualitative analysis. The qualitative evaluation can give valuable information that can complement the quantitative data. It is used as triangulation and the crosschecks of the results.

3. PILOT STUDY

The pilot study and the usability testing of the VR programs were implemented at the same time. There are three aims in the pilot study:

1. To obtain preliminary data for evaluating the effectiveness of VR programs and psychoeducational group in training street survival skills of people with stroke
2. To compare the effectiveness among groups
3. To test the usability of newly developed VR programs in training street survival skills

Figure 1. *Screenshots from the VR programs*

The pilot study has recruited 11 subjects in total and randomly allocated them into 3 groups: VR group psychoeducational group and control. They met the criteria used in the main study. The training contents were the same in two intervention groups but different in the media. In VR group, the subjects went through 6 sessions of 1-hour training in street survival skills with the newly developed VR programs. These programs were installed in a desktop personal computer (PC) with 17" monitor, keyboard, mouse and joystick. The subjects could choose their preferred input devices for navigating and interacting within the virtual environment so that they had more self-directed actions during the training. The PCs were placed in a quiet and comfortable room to minimize the distractions so that the subjects can concentrate on the training.

For the program development, these programs were based on real life photo in Hong Kong and compiled by using LivePicture and C++ programming to create the virtual environment. Two major tasks in the virtual environment were crossing road safely and using underground railway (MTR). The VR programs compiled by LivePicture could provide the overall flow in the virtual environment. The virtual journey started at home and required the subjects go to a destination at the end. Inside the virtual environment, the subjects could freely navigate through the panorama (a 360 degree scene) and look around the environment. They were required to walk along the streets to find out the suitable place to cross the road safely and identify the MTR entry. After going down to the concourse, the subjects should buy tickets and pass through the gate. Then went down to the MTR platform where the train will come. They would take on the MTR and arrive at a designated station where they would take off. Then they would get out from the MTR station through the concourse. Finally, the subjects had to walk along the streets and find the way to arrive at the entrance of a building. Besides this VR program, the others compiled by C++ programming were also developed to allow more specific training in two major tasks. The subjects could control a virtual person to practice crossing road and using MTR separately. They could do more delicate motions within that virtual environment. The first session was to introduce the learning aims, methods and contents as well as provide the time for the subjects to get familiar of the user interfaces and input devices. The baseline assessment had been conducted in this session. The second and third session allowed the subjects freely explore and navigate within the virtual environment. Also, the therapists have discussed with the subjects for the possible problems and try to work out the potential solutions. The forth and fifth session required them to perform several specific actions as a task-specific training. The post-treatment assessment in real life situation was done for outcome measures in the sixth session.

In psychoeducation group, the training contents were the same as the VR group. The PowerPoint would be presented to the subjects so that they could learn the ideas of the skills. Psychoeducation approaches would be adopted in which demonstration (photo, video or therapist), role-play practices, coaching, feedback and homework assignment were provided to facilitate the learning of the skills. The arrangement of the sessions were similar of those of VR group that started with baseline assessment and ended up with post-treatment assessments. In control group, only the pre and post assessments have been done.

In the pilot, the independent variable was the different groups in the persons with stroke as well as other demographic information. The dependent variables were the subjects' knowledge and skills and their self-efficacy in applying those learned skills in their daily functioning. They were in terms of the results of the self-efficacy scale, behavioural checklist and the formative evaluation. Descriptive statistics have been used to describe the demographic data and the performance at different time and non-parametric tests have been used to analyze the performance within and among groups.

Since the VR programs were newly developed for the research, the information about their usability was almost unknown. It was necessary to conduct the usability testing to investigate in this aspect and the insights gained could help to improve the VR programs so that the potential clients could use and learn the programs easier. As the clients become more motivated in the VR programs, the training results should be easier to be effective and therapeutic. Usability is a quality attribute that assesses how easy user interfaces are to use. It also refers to methods for improving easy-to-use and easy-to-learn during the design process (Nielsen, 2003). The usability testing could provide substantial improvements in the program design and it is especially effective in the early stage of development. 5 subjects were enough for the users testing. The number of usability problems found in a usability test with n users is: $N(1-(1-L)^n)$, where N is the total number of usability problems in the design and L is the proportion of usability problems discovered while testing a single user. The typical value of L is 31%, averaged across a large number of projects they studied (Nielsen, 2003). With a single test user, researchers learn almost a third of all there is to know about the usability of the design. As adding more and more users, new information would become less and less. After the fifth user, it is a waste of time by observing the same findings repeatedly but not learning much new (Nielsen, 2003). During the usability testing, the subjects were observed while they were actually using the VR programs. Special notes was made on what they succeed, what they fail, how they deal with the problems and so on. After the testing, the subjects were interviewed in a semi-structured manner in which special issues or constructive feedback could be solicited. The data were tabulated and analyzed in terms of brief description of the task, usability question, quantitative goal, quantitative goal achieved, descriptive statistic, problems identified and suggested solutions. The results were used to refine the 2DVR programs that will be used in the main study.

4. RESULTS

11 subjects, 6 males and 5 females with mean age of 46.5±5.0 and mean suffered year of 2.2±0.6, were recruited in the pilot. 5 were left hemiplegia and 6 were right hemiplegia. Most of them were ADL independent with secondary school education level. 3 knew very little about computer and others have basic to moderate computer knowledge. All of them were mobility independent with adequate functions in hearing, vision, communication and comprehension. They were randomly allocated in 3 groups and resulted in 3 in control, 3 in psychoeducation and 5 in VR group. For control group, pre self-efficacy mean was 98.3±3.5; post self-efficacy mean was 98.3±3.5; pre behavioural skills was 58.0±2.6; post behavioural skills was 58.0±2.6. For psychoeducation group, pre self-efficacy mean was 97.0±4.4; post self-efficacy mean was

102.0±8.9; pre behavioural skills was 57.0±1.7; post behavioural skills was 59.7±0.6. For VR group, pre self-efficacy mean was 99.0±8.4; post self-efficacy mean was 103.0±5.7; pre behavioural skills was 53.6±8.0; post behavioural skills was 64.8±1.8.

Due to small sample size, non-parametric Kruskal-Wallis test was used to compare among groups. It had shown that pre self-efficacy, post self-efficacy and pre behavioural skills had no significant difference among groups but post behavioural skills had the significant difference among groups (Chi-Square=8.073, df=2, p=.018). It meant that the groups were homogeneous prior to the treatment and the substantial difference in post behavioural skills was found after treatment. To further investigate the difference between groups, Mann-Whitney U test was performed and the significant differences in post behavioral skills were found in control vs VR (U=.000, Z=-2.291, p=.022) and psychoeducation vs VR (U=.000, Z=-2.306, p=.021). It meant that VR could result in higher behavioral skills. Wilcoxon signed ranks test was used to compare pre and post measurements within groups. Both control and psychoeducation group had no significant difference in pre-post comparisons of self-efficacy and behavioral skills. Only VR group has shown significant difference in pre-post behavioral skills comparisons (Z=-2.032, p=.042) while the self-efficacy were not significant (Z=-1.826, p=.068). It meant that VR could improve the behavioral skills.

All 5 subjects in VR group have also performed the usability testing and the descriptive statistics were conducted. A number of usability tasks were requested to perform to test the user interface and the control devices. The tasks included moving with keyboards, changing view with mouse, following model, mouse actions and zooming. Accomplished times were measured with stopwatch from the instruction announced to the finished tasks. Generally, the subjects could do the tasks within reasonable time. Zooming involved coordinated use of mouse and keyboard simultaneously, so it could be harder and the accomplished times were longer.

Table 1. *Descriptive statistics of usability tasks*

Tasks	Move U/D	Move L/R	View U/D	View L/R	Follow model	Mouse click	Mouse drag	Mouse release	Mouse move	Zoom in	Zoom out
Min	2.00	2.00	3.00	3.00	3.00	2.00	3.00	1.00	3.00	3.00	3.00
Max	3.00	3.00	4.00	4.00	5.00	3.00	5.00	3.00	5.00	7.00	7.00
Mean	2.6000	2.6000	3.4000	3.4000	3.6000	2.2000	3.8000	1.6000	4.0000	4.8000	4.8000
SD	.5477	.5477	.5477	.5477	.8944	.4472	.8367	.8944	.7071	1.4832	1.4832

The common feedbacks from the subjects about use of VR programs in training were interesting, realistic, funny and new. Most of them found the control was easy but some of them wanted a third person view instead of the first person view used in the pilot study. They thought that it would be easier to judge the distance. After the treatment, most of them felt a bit tired but no motion sickness or other side effects were reported. Some of them said that the VR programs could make them became attentive to signs in real life situation and familiarized the procedures of ticket machines. Most of them wanted the expansions in VR programs such as more locations or buildings, more transportation methods and map readings. A few of them wanted harder tasks such as banking, shopping and sports.

5. DISCUSSIONS

Because of small sample size in the pilot, the power of study may not be adequate to show the significant difference among groups and within group. Only VR group shown significant difference in behavioural skills during group and pre-post comparisons. The main study with larger small size and insights from pilot should yield more fruitful results. By the way, transfer of skills from a virtual environment to the real world is possible and it helps cognitively impaired individuals relearn important daily living skills (Burdea et al, 2000; Riva et al, 2000; Riva, 1998). Moreover, immerse virtual kitchen was developed in which patients with TBI can perform a meal preparation task involving 30 steps (Christiansen et al, 1998; Zhang et al, 2001). Persons with TBI consistently demonstrated a significant decrease in the ability to process information, identify logical sequencing and complete the overall assessment, compared with persons without TBI (Zhang et al, 2001). The acceptable reliability and validity supports further development of virtual environment as an assessment and treatment tool.

6. CONCLUSIONS

In the near future, more researches will be conducted for investigating the use of VR in rehabilitation since VR is an effective tool for rehabilitation, with its own superiority. The computer graphics are better now and the 3D rendering techniques become more mature, thus contributing to the reality. The software programming has all-round abilities thus contributing to the versatility. The other advantages include that VR can provide a controlled and structured environment. It provides unlimited opportunities for acquiring and processing information, exploring and learning. The patients can do the tasks repeatedly under little supervision. Immediate feedback, prompts, cues in various sensory modalities can also be incorporated into VR so that it can reinforce the desirable response and motivate the patients to continue rehabilitation. Moreover, VR offers a safe environment with little threats so that some barely possible practices in the real world can be done within VR before actual performance. Modifiability and customizability can provide a tailor-made, optimal learning situation for each patient, thus enhancing the generalization of the learning. The advantage of recording each action can be used for further retrieval, analysis and documentation of the patient's progress. For these reasons, virtual reality should be a suitable and effective tool for the stroke rehabilitation.

7. REFERENCES

A Bandura (1997), *Self-efficacy: the exercise of control*. New York: W.H. Freeman

G Burdea, V Popescu, V Hentz et al. (2000),Virtual reality-based orthopedic telerehabilitation. *IEEE Transactions on Rehabil Engine*; 8(3):430-432

C Christiansen, B Abreu, K Ottenbacher, et al. (1998), Task performance in virtual environments used for cognitive rehabilitation after traumatic brain injury. *Arch Phys Med Rehabil*;79:888-92

TY Chuang, WS Huang, SC Chiang, et al. (2002), A virtual reality-based system for hand function analysis. *Computer Methods & Programs in Biomed*;69:189-196.

ST Gontkovsky, NB McDonald, PG Clark, WD Ruwe (2002), Current directions in computer-assisted cognitive rehabilitation. *Neurorehabilitation*, 17, 195-199

MA Grealy, DA Johnso, SK Rushton (1999), Improving cognitive function after brain injury: the use of exercise and virtual reality. *Arch Phys Med Rehabil*;80:661-7

HG Hoffman, DR Patterson, GJ Carrougher, et al. (2001), Effectiveness of virtual reality-based pain control with multiple treatments. *Clin J Pain*;17(3):229-235

HG Hoffman, DR Patterson, GJ Carrougher (2000), Use of virtual reality for adjunctive treatment of adult burn pain during physical therapy: a controlled study. *Clin J Pain*;16(3):244-250

J Nielsen (2003), Usability 101, *http://www.useit.com/alertbox/20030825.html*

J Nielsen (2003), Misconception about usability, *http://www.useit.com/alertbox/20030908.html*

G Riva, L Gamberinic (2000), Virtual reality as telemedicine tool: technology, ergonomics and actual applications. *Tech & Health Care*;8:113-127

G Riva (1998), Virtual environments in neuroscience. *IEEE Transactions on Infor Tech Biomedi*; 2(4):275-281

AA Rizzo, JG Buckwalter, U Neumann. (1997), Virtual reality and cognitive rehabilitation: a brief review of the future. *J Head Trauma Rehabil*;12(6):1-15

L Zhang, BC Abreu, B Masel, et al. (2001), Virtual reality in the assessment of selected cognitive function after brain injury. *Am J Phys Med Rehabil*;80:597-604.

Robot aided therapy: challenges ahead for upper limb stroke rehabilitation

R C V Loureiro, C F Collin and W S Harwin

Department of Cybernetics, The University of Reading,
Whiteknights, Reading, RG6 6AY, UK

r.c.v.loureiro@reading.ac.uk, Christine.Collin@rbbh-tr.anglox.nhs.uk, w.s.harwin@reading.ac.uk

www.isrg.reading.ac.uk

ABSTRACT

People who have been discharged from hospital following a stroke still have a potential to continue their recovery by doing therapy at home. Unfortunately it is difficult to exercise a stroke affected arm correctly and many people simply resort to using their good arm for most activities. This strategy makes many tasks difficult and any tasks requiring two hands become nearly impossible. The use of haptic interface technologies will allow the reach and grasp movements to be retrained by either assisting movement, or directing movement towards a specified target. This paper demonstrates how initial work on machine mediated therapies can be made available to a person recovering at home.

1. INTRODUCTION

Stroke is a leading cause of disability in the UK, with incidence rates between 1.25 and 1.8 per 1000 people per annum with the rate higher in Scotland and higher for men (Stewart et al, 1999; Scottish Health Statistics, 2002). Traditional treatments rely on the use of physiotherapy that is partially based on theories and also heavily reliant on the therapists training and past experience. The lack of evidence to prove that one treatment is more effective than any other makes the rehabilitation of stroke patients a difficult task. Upper limb motor re-learning and recovery levels improve with intensive physiotherapy. The need for conclusive evidence supporting one method over the other and the need to stimulate the stroke patient clearly suggest that traditional methods lack high motivational content, as well as objective standardised analytical methods for evaluating a patient's performance and assessment of therapy effectiveness. Although the causes of stroke are well known and it is possible to reduce these risks, there is still a need to improve rehabilitation techniques.

2. CURRENT STATE-OF-THE-ART

2.1 Early interventions

According to physiotherapy literature, attention and motivation are key factors for motor relearning following stroke (Yekutiel, 2000). One way to achieve early intensive interventions is via machine mediated therapies but there is a lack of good tools available to the therapist.

Several authors have already proposed the use of robotics for the delivery of this type of physiotherapy. The first far-reaching study on acceptance of robot technology in occupational therapy for both patients and therapists was done by Dijkers and colleagues using a simple therapy robot (Dijkers et al, 1991). Dijkers study reports a wide acceptance from both groups, together with a large number of valuable suggestions for improvements. Advantages of Dijkers therapy include the availability of the robot to successively repeat movements without grievance, as well as, the ability to record movements. However, there was no measure of movement quality and patient cooperation was not monitored.

2.2 Recent research

More recent studies however, follow a more task oriented approach. Johnson et al, (1999) have developed the SEAT: "simulation environment for arm therapy" to test the principle of the 'mirrored-image' by the provision of bimanual, patient controlled therapeutic exercise. The device comprises of a customised design of a car steering wheel equipped with sensors to measure the forces applied by patient's limbs, and an electrical motor to provide pre-programmed assistance and resistance torques to the wheel. Visual cues where given to the patient via a PC-based driving simulator that provided graphical road scenes. The interface allowed the participation of the patient in the task and the involvement of the paretic limbs in the exercise (Johnson et al, 2003).

Based on the same mirror image concept, Lum and colleagues (Lum et al, 1999) at VA Palo Alto research introduced the MIME: "Mirror-image motion enabler". The initial MIME prototype used a Puma-260 robot coupled through a force and torque transducer to one of the forearm splints used to support the patients arms weigh. The splints were free to rotate and tilt at the end of modified mobile arm supports. In the current MIME workstation, the robot is a Puma-560, the paretic limb mobile arm support is eliminated, and a 6DOF position digitiser replaces the contra-lateral support. The Puma-560 facilitates unilateral therapeutic exercises in 3 modes and 12 trajectories. A computer controls movement of the robot, with specific pre-programmed tasks tailored to the subject's level of recovery and therapeutic goals. Clinical trials with 27 chronic stroke patients (> 6 months post stroke) based on Fugl-Meyer exam, have shown that low compliance systems do not influence negatively the upper limb joint passive range of motion and pain. Results also suggested that robot-aided therapies are safe and effective for neuro development treatment (Shor et al, 2001). This work is now being commercialised by Dr. Mahoney at Applied Resources, USA.

Work done at MIT by Krebs et al, (1999) on the development of a new robot that allowed the patient to exercise against therapist nominated stiffness and damping parameters uses a different approach from the systems described so far and is the project with more exposure to patients in the literature to date. Their MIT-MANUS a 3DOF (2DOF active 1DOF passive) planar manipulator, has performed a series of clinical trials since 1995 at Burke Rehabilitation Hospital. Recent results were reported (Krebs et al, 2001) from a total of 76 patients assessed for upper limb subsection of the Fugl-Meyer test, motor power for shoulder and elbow, motor status score for shoulder and elbow, and motor status score for wrist and fingers. It was shown that the manipulation of the impaired limb influences recovery, the improved outcome was sustained after 3 years, the neuro-recovery process continued beyond the commonly accepted 3 months post-stroke interval, and the neuro-recovery was dependent on the lesion location. The MIT-MANUS mechanism however limits the range of possible therapies, has limited data collection facilities and does not allow bimanual therapies as the SEAT and MIME systems.

2.3 Gentle/s project

Gentle/s was a three year project funded by the European Commission to develop machine mediated therapies for neurorehabilitation of people with stroke. This project resulted in 3 prototype machines (Harwin et al, 2001). Gentle/s had the goal to both improve quality of treatment and reduce costs. A pilot study showed that subjects were motivated to exercise for longer periods of time when using a mixture of haptic and virtual reality systems (Loureiro et al, 2003).

Subjects used the Gentle/s prototype while seated at a table (figure 1). The person's arm was put in an elbow orthosis with wires suspending it from an overhead frame so as to eliminate the effects of gravity and address the problem of shoulder subluxation. Software was developed to manage the data collection, control and simulate complex virtual worlds (Loureiro et al, 2003). Subjects using the Gentle/s system could exercise "reach-and-grasp" type of movements (without the grasp component) in 3 degrees of freedom through interaction with a virtual room (figure 2 and 3). Task oriented activities can easily be created by a therapist using a graphical user interface and visual guidance is provided in the form of start/end points for a specific movement pattern. Control of human movement is based on minimum jerk theory and on a novel methodology that uses a spring and a damper combination that moves on a constrained path, called "bead pathway" (Amirabdollahian et al, 2002). The software also provides 3 different levels of movement control. The first called "patient passive mode" is aimed at users on the initial phases after the stroke, where the haptic device will teach the correct movements. The second, called "active-assisted mode" helps the user to complete a determined movement. The third mode, "active mode" provides only correction of movement. Clinical trials (Amirabdollahian et al, 2003; Coote et al, 2003) were conducted at Trinity College, Dublin and at Battle Hospital, Reading (total of 30 patients) and show the system was effective in improving recovery and acceptable to both subjects and therapists. The data from the principal study is in accordance with findings of the MIT MANUS studies (Krebs et al, 2001). One of the conclusions to take from the Gentle/S study is that this type of therapy appears to be particularly appropriate for more severely disabled patients. Further studies are needed to establish the level, nature, onset and determination of treatments that will result in the best recovery for an individual patient.

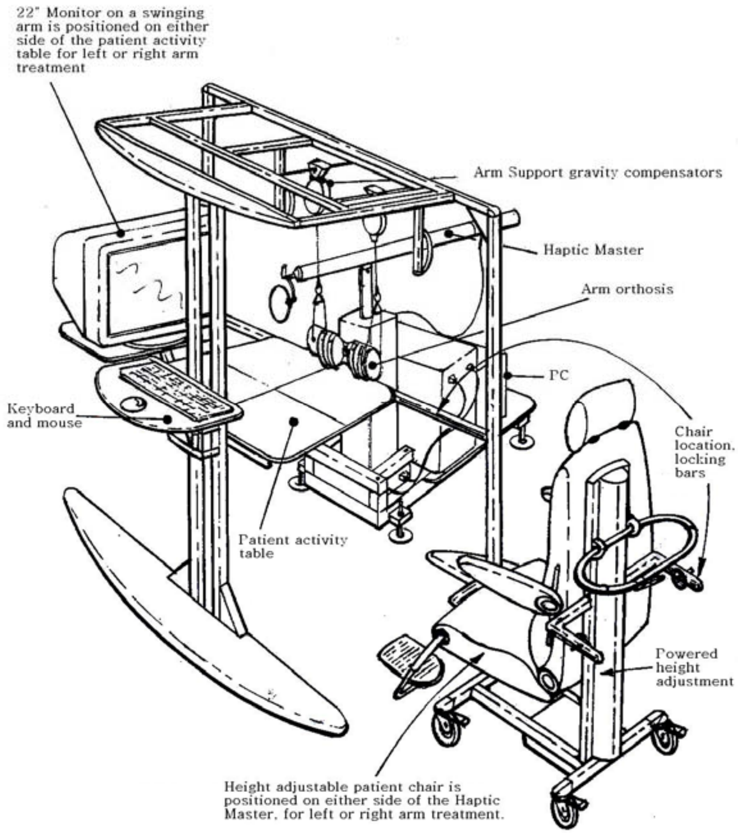


Figure 1. *The Gentle/s prototype system, showing the overall frame, an elbow orthosis suspended from the overhead frame, the HapticMaster with gimbals, the exercise table, computer screen and the wheelchair.*

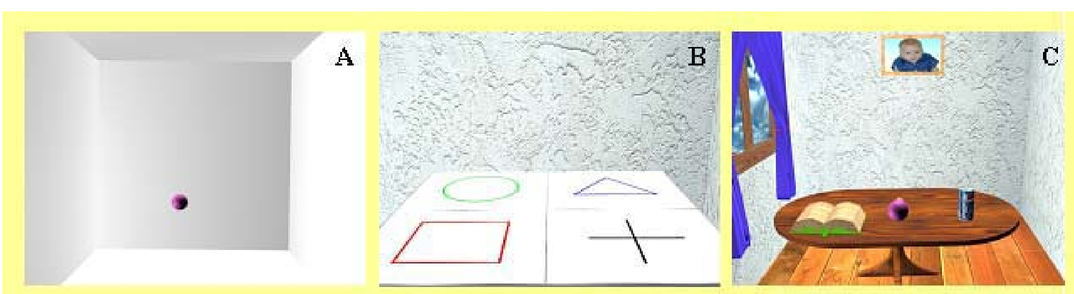


Figure 2. *Some of the virtual rooms available in the Gentle/s system. A) a simple environment that represents the haptic interface workspace and intends to provide early post-stroke subjects with awareness of physical space and movement; B) an environment that resembles what the patient sees on the table in the real world. This environment was developed to help discriminating the third dimension that is represented on the computer 2D screen; 3) A high detail 3D graphical environment of a room aimed at providing more motivational exercise activity.*

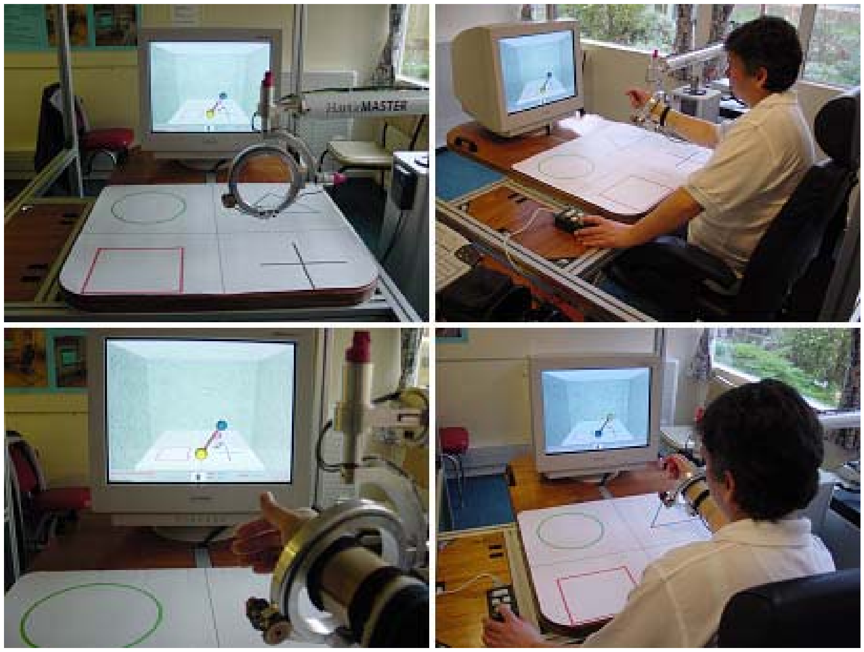


Figure 3. *Subject using the Gentle/s prototype system with one of the available virtual rooms. The virtual environment replicates the user's viewpoint in the real world where the mat with shapes is used to help perception of movement and space.*

3. THE CHALLENGE I: TECHNOLOGY NEEDED

The technologies presented on the previous section have the potential to revolutionise the way hospitals operate and hence, reduce the cost of stroke rehabilitation by allowing therapists and clinicians to manage a larger number of patients in the same amount of time. It is also likely that the recovery process can be reduced if this type of therapies is administrated at the correct intensity and method, at the acute phase of stroke. More clinical evaluation of current technologies is still needed to evaluate this type of approach, by increasing the number of subjects exposed to this type of therapies and to make the technology widely available, from its usage in hospitals to the patient's home. Gentle/s has identified the need of using the hand as well as the forearm and upper arm in the rehabilitation process of the upper limb using VR. Exposure of Gentle/s to the media (BBC, Discovery channel) generated a high level of interest in home based systems. This interest from people with a stroke, coupled with the results from the Gentle/s project, are a strong indication of a need for a high quality home therapy approach to stroke rehabilitation.

Reinkensmeyer has demonstrated home based rehabilitation using an adapted 'force feedback' games joystick (Reinkensmeyer et al, 2001). Therapies were downloaded from a web site and these could provide physically assist or resistance to movement as the user exercises. The exercises were quantified giving a level of feedback of performance, thus allowing users and their caregivers to assess rehabilitation progress. The study was small and suffered from the very small workspace and low level of force assistance provided by the joystick but it gives a good indication of the acceptability of this approach.

It is clear from our work in the Gentle/s project that there is a need to develop the technology in two areas. The relatively low levels of media exposure showed the need for a device that can be loaned to well motivated patients on leaving inpatient rehabilitation services. This device would be 'on loan' to the patient during possibly the 6 months following discharge and would provide challenging and motivating therapies to the individual, and at the same time, communicate with a central server to report usage, and update treatments. It is apparent that such a machine needs to have a relatively low cost, but more importantly needs to fit within the person's home environment so qualities such as light weight, small foot print and low power consumption will necessarily limit the abilities of the technology. In this regards the requirement of the Gentle/s technologies to enforce (albeit with varying levels of encouragement) a particular movement will probably need to be relaxed. A system based on guided weightless arm support is a better alternative as it can better meet the needs of a home based system. It may well be possible to use the person's non-affected hand to provide guidance by developing the ideas championed by Lum et al, (1999) in the MIME approach. It would be relatively straightforward to provide the car steering mechanism suggested by Johnson et al, (2003) as a home based system although this single application area may be too restrictive to promote reasonable recovery of motor movements in a large workspace.

The need for full functioning hospital systems remains, although considerably more work is needed to assess both the usage and the impact of this type of device in a hospital setting. There is a need for better assessment methods in stroke rehabilitation, separated from the intervention. Simplistic and subjective scores collected over a short time period are the norm for most assessments and these are insensitive to functional recovery. Furthermore, there is a strong need to move this type of treatment into the acute phase of stroke recovery, ideally beginning machine mediated therapies within a few days of admission. In many places this requires a hospital ward based machine that, although under the jurisdiction of the therapy staff, can be set up and run by the nursing staff as per the therapist's prescription for the patient (in the same way drugs are administered by nursing staff to the doctors prescription). This machine would need all the technical aspects of the Gentle/s system of having a reach commensurate with upper limb movement, requiring and encouraging patient involvement and giving goal directed therapies. It also should promote reach & grasp, manipulation and a combination of reach grasp and manipulation, in contrast to the Gentle/s prototype that only focused on reaching movements.

4. THE CHALLENGE II: QUESTIONS TO BE ANSWERED

Research on the area of rehabilitation robotics and recent results obtained with the Gentle/s system raised several research issues in the perspective of stroke, such as the need to be able to rehabilitate both reach and grasp in a context that allows the manipulation of objects in virtual tasks. It is not clear if recovery levels are attributed to the haptics alone or to a combination of visual and haptic cues, and performance cues need to be investigated further.

The hypothesis is that by giving individuals who are recovering from a first stroke access to motivating and challenging therapy 'on demand' and at their convenience at home, they will exercise for longer in a physiologically appropriate way that will lower their level of impairment and their disabilities. These therapies must be delivered in a structured environment that is well monitored so they are acceptable to the clinicians responsible for the management of the patient's recovery.

Arm movements are so common that we easily forget how complex they are. Even a simple movement like reaching for a can of soft drink requires fine-tuned muscle activation patterns. At present the authors are working with the aim to derive engineering based conclusions about stroke by the investigation of the effects of haptics and virtual environments in the rehabilitation process of the upper limb in stroke rehabilitation of the hand, from a systems point of view in order to identify better treatment alternatives.

The authors consider that better functional and motor recovery outcomes of stroke patients will be achieved where patients receive a repetitive, challenging and motivational machine mediated therapy that encourage reach and grasp movements. Part of our current and future work is devoted to the development of modular systems (arm support mechanisms, grasp assistance modules, arm reach movement assisting modules based on haptic interface technology, movement control software, data management, etc) that can be placed on a home context. Arm support mechanisms can be in the form of a forearm orthosis that can replicate the support mechanism developed for the Gentle/s project but instead of a large frame, it could be mounted on a chair. Grasp assistance modules could be in the form of a mechanism that can either simulate the force/position response of a grasp, or where the person has no discernable ability to move their fingers, move the fingers and thumb in a pattern that mimics the natural grasp of an object. As for arm reach movement assisting modules, these would assist with the reach phase of a reach and grasp movement. This can be either by constraining the person's arm to the 'correct' path, or if the person can not make the complete movement, the arm assistance module would move the persons arm along a nominally 'correct' path.

It is equally important to develop models of human control that are clinically useful and mechanically accurate. For example, if upper limb is subject to mechanical loading for a period of time, can we determine if the brain will over time adapt and adjust to added stiffness, mass and inertia? One can then postulate that stroke magnitude can be reduced if subjects are submitted to training in a context similar to the Gentle/s system. There is a need to understand stroke from a systems perspective, that is model the response to mechanical loading and hence develop better therapeutic and restorative methodologies. These and other questions are the source of our motivation to contribute to the development of the field of machine mediated neurorehabilitation to amass evidence on the clinical viability and the commercial potential of modular home based systems. Preliminary results on this area will be of interest to both clinical and commercial players in this field and will enable the larger studies that are needed for bodies such as NICE (National Institute for

Clinical Excellence). The long term purpose is to introduce revolutionary thinking into the traditionally under-regarded field of rehabilitation.

Our drive is also reflected on the challenges of reducing the current costs of systems such as Gentle/s that can be used at home and focusing on the rehabilitation of the entire upper limb. This includes the need to investigate; 1) real and virtual grasp exercises with healthy subjects to aid the development of linear and parametric models characterising the dynamic properties of the hand, wrist and forearm of stroke victims. 2) user involvement for increased recovery levels and this is reproduced by the level of interaction and the feedback provided to the user. Such level of interaction in virtual tasks need to be enhanced by means of the addition of active hand orthoses (exoskeleton) to the current systems, which should allow for grasping, monitor and parameterisation of the hand for different types of grasping activities. Bearing in mind that, physiotherapy practice is not guided explicitly on theories or research on the literature but on the experience that therapists acquire by working with patients and other experts in the field (and obviously on the type of approach that they were trained on) it is paramount that assessment methods become more objective. Traditional physiotherapy assessment methods are inconsistent varying from one therapist to another and from hospital to hospital, consequently it is important that the new machine mediated therapies are able to provide useful objective measures of patients' performance that can be easily analysed by clinicians and therapists alike. Similarly, motivational techniques should be taken into consideration when developing activities based on virtual environments and therapy performance cues in the form of encouraging feedback (sounds, visual stimulation, and task completion reward schemes) should be provided to the user.

5. CONCLUSIONS

In the first part of this paper we present the current state-of-the-art on machine mediated therapies for upper limb stroke rehabilitation, where a review of past/present research studies in this area is discussed and the Gentle/s prototype placed in perspective within the available technology.

The second part concentrates on our research commitments and on our views of what lies ahead. Our motivation is towards innovative robotic technology that could be placed into a home package to allow patients with upper limb weakness due to stroke, to practice meaningful interactive exercises in a virtual environment, with robotic delivered feedback and guidance. We have already shown that machine mediated therapies (MMT) can be delivered in the hospital environment to stroke patients with upper limb weakness with no harmful side effects and with positive results for some individual patients. Similarly, this work has been replicated elsewhere.

There is now a need to test whether machine mediated therapies can be effectively delivered in a home environment, thus benefiting from the cosy, relaxed home environment to reduce impairment and disability, and therefore reducing the rehabilitation associated costs. Low cost modular home based systems, can also provide patients with the opportunity to work as hard as they want to on a particular disability. This approach will also allow the patient the opportunity to demonstrate initiative and independence in their individualised home-based rehabilitation programme.

Acknowledgements: The first author acknowledges the financial support from the Portuguese Ministry of Science and Higher Education, Fundação para a Ciência e a Tecnologia (under research grant SFRH/BD/12271/2003).

6. REFERENCES

DJ Reinkensmeyer, CT Pang, JA Nessler, and CC Painter (2001), Java therapy: web-based robotic rehabilitation *Proc. Int. Conf. Rehab. Robotics, Integration of Assistive Technology in the Information Age,* M. Mokhtari (Ed.), IOS Press, pp.66-71.

F Amirabdollahian, R. Loureiro, and W. Harwin (2002), Minimum Jerk Trajectory Control for Rehabilitation and Haptic Applications, *Proc. IEEE Intl Conf. on Robotics & Automation,* Washington, DC, pp.3380-3385.

F Amirabdollahian, E. Gradwell, R. Loureiro, C. Collin, and W. Harwin (2003), Effects of the Gentle/s Robot Mediated Therapy on the Outcome of Upper Limb Rehabilitation Post-Stroke: Analysis of the Battle Hospital Data. *Proc. of 8th International Conference on Rehabilitation Robotics,* Published by HWRS-ERC Human-friendly Welfare Robot System Engineering Research Center KAIST, Republic of Korea, pp.55-58.

HI Krebs, N Hogan, N., BT Volpe, ML Aisen, L Edelstein, and C Diels (1999), Robot-Aided neuro-Rehabilitation in Stroke: Three-Year Follow-up, *IEEE Intl. Conf. Rehabilitation Robotics,* Stanford, CA, USA.

HI Krebs, BT Volpe, J Palazzolo, B Rohrer, M Ferraro, S Fasoli, L Edelstein, and N Hogan (2001), Robot-aided neuro-rehabilitation in stroke: interim results on the follow-up of 76 patients and on movement performance indices, *Proc. Int. Conf. Rehabilitation Robotics, Integration of Assistive Technology in the Information Age,* Mounir Mokhtari (Ed.), IOS Press, 9, pp. 45-59.

JA Stewart, R Dundas, RS Howard, AG Rudd and CDA Wolfe (1999), Ethnic differences in incidence of stroke: prospective study with stroke register, *British Medical Journal (BMJ),* 318*, pp. 967 – 971.*

M Yekutiel (2000), *Sensory re-education of the hand after stroke,* Whurr Publishers, London.

MP Dijkers, PC deBear, RF, Erlandson, K Kristy, DM Geer, and A Nichols (1991), Patient and staff acceptance of robot technology in occupational therapy: a pilot study, *Journal of Rehabilitation Research and Development,* 28, *2*, *pp. 33-44.*

MJ Johnson, HFM Van der Loos, CG Burgar, and LJ Leifer (1999), Driver's SEAT: simulation environment for arm therapy, *Proc. IEEE Intl. Conf. Rehabilitation Robotics,* Stanford, CA, USA, pp. 227-234.

MJ Johnson, HFM Van der Loos, CG Burgar, PC Shor and LJ Leifer (2003), Design and evaluation of Driver's SEAT: A car steering simulation environment for upper limb stroke therapy, *Robotica,* **21,** *pp.13-23.*

PC Shor, PS Lum, CG Burgar, HFM Van der Loos, M Majmundar, and R Yap (2001), The effect of robotic-aided therapy on upper extre,ity joint passive range of motion pain, *Proc. Int. Conf. Rehabilitation Robotics, Integration of Assistive Technology in the Information Age,* Mounir Mokhtari (Ed.), IOS Press, 9, pp. 79-83.

PS Lum, CG Burgar, DE Kenney, and HFM Van der Loos (1999), Quantification of force abnormalities during passive and active-assisted upper-limb reaching movements in post-stroke hemiparesis, *IEEE Transactions on Biomedical Engineering,* 46, *6, pp. 652-662.*

R Loureiro, F. Amirabdollahian, M. Topping, B. Driessen, and W. Harwin (2003), Upper Limb Mediated Stroke Therapy - GENTLE/s Approach, *Special Issue on Rehabilitation Robotics Journal of Autonomous Robots*, Kluwer Academic Publishers, 15, *1, pp. 35-51 (ISSN 0929-5593).*

Scottish Health Statistics (2002), *ISD Scotland National statistics release,* web resource: http://www.isdscotland.org/isd/info3.jsp?pContentID=2857&p_applic=CCC&p_service=Content.show& *(Accessed on 14-07-04)*

S Coote, EK Stokes, BT Murphy, WS Harwin (2003), The effect of GENTLE/s robot mediated therapy on upper extremity function post stroke, *Proc. of 8th International Conference on Rehabilitation Robotics,* Published by HWRS-ERC Human-friendly Welfare Robot System Engineering Research Center KAIST, Republic of Korea, pp.59-61.

W. Harwin, R. Loureiro, F. Amirabdollahian, M. Taylor, G. Johnson, E. Stokes, S. Coote, M. Topping, C. Collin, S. Tamparis, J. Kontoulis, M.Munih, P. Hawkins, and B. Driessen (2001), The Gentle/s project: a new method for delivering neurorehabilitation, *Assistive technology - added value to the quality of life AAATE'01,* IOS Press Amsterdam C. Marincek et al. pp.36-41.

ICDVRAT 2004

Session II. Training and Virtual Transport

Chair: Hideyuki Sawada

HUMANICS 1 – a feasibility study to create a home internet based telehealth product to supplement acquired brain injury therapy

A Lewis-Brooks

Computer Science Department, Aalborg University Esbjerg,
Niels Bohrs vej 8, Esbjerg, DENMARK

tonybrooks@cs.aue.auc.dk

www.cs.aue.auc.dk, www.cfh.ku.dk/6_forskning/Humanics

ABSTRACT

The goal to produce a unique, cost effective, and user-friendly computer based telehealth system product which had longevity and the ability to be integrated modularly into a future internet-based health care communication provision was conceptualised as an aid to home-based self-training. This through motivated creativity with the manipulation of multimedia. The system was to be a supplementary tool for therapists. The targeted group was initially to be those with acquired brain injury. This paper details phase 1 of the product feasibility testing.

1. INTRODUCTION

Acquired brain injury rehabilitation entails a long and enduring process towards training the individual to a realisation of potentials so as to be able to live a life with optimal quality following injury. Training often involves travel to a clinic which involves certain stressful situations, economic considerations as well as environmental consequences. A system design where private individuals could be motivated to train at home and utilizing the internet send their progress information to the clinic therapist for management was novel and judicious. Furthermore the all essential support from family members could provide added motivation as all are capable of 'playing' together with the system in the home setting. The question was asked however whether a generic system could be created that would be 'user-friendly' and efficient across age groups, ability groups and have continued worth over novelty value as is often seen in similar 'tools'. This paper chronicles the first phase of the research which was established as a feasibility study to ascertain if the target group could benefit from such a product and to receive their input. In so doing it lays out the foundation and philosophies involved.

2. BACKGROUND

At the 'Year of the Brain' conference in Aalborg, Denmark, the author presented his 'SoundScapes' body of work (Brooks 2003). Prominent figures from two of the leading Centres for Rehabilitation of Brain Injury were in the audience and, understanding the potentials inherent in the concept, they approached the author which resulted in trial sessions being initiated with physiotherapists and acquired brain injured patients. The sessions led to further collaborations and presentations including one where the author won the top European prize at the BAM (Brokerage Applied Multimedia) event hosted by the Eureka organisation in Stockholm, Sweden. The presentation was of his evolving research within the field of special needs and his design of the proposed telehealth system product which is the subject of this paper. Subsequently in a collaborative effort with the same team from the Centre for Rehabilitation of Brain Injury in Copenhagen the project which is covered by this paper was initiated which resulted in a Danish government funding of the study.

The Humanics study that this paper refers to was covered by non-disclosure agreements under the Danish government product development contract until recently (2004) whereupon written clearance was given by the CRBI enabling the author to publish on the project.

3. CENTRE FOR REHABILITATION OF BRAIN INJURY (CRBI)

The CRBI was established in Copenhagen, Denmark in 1985 for the rehabilitation of acquired brain injury through holistic & individual treatment with a main focus on return to work and/or improved quality of life. Today, in 2004, it is a self contained entity based at Copenhagen University following in 1993 being awarded the Special Institute Status and accordingly it initiates training, research and education programs. It is one of the top European Centres and the team consisting of Psychologists, Physiotherapists, Neurologists and Speech Therapists is highly respected in the field. The aim of CRBI is stated as:

Integration through:-
'Training and awareness of psychological & physical deficits & strengths' and 'Insight & Compensation techniques'

4. PROJECT CONCEPT

Most people joyously appreciate moving or dancing to music. *Creating* music and images with body movement alone adds a new and different dimension to such joy. Physical rehabilitation after sustained brain injury is often an enduring and cumbersome task to the patient, who is only encouraged if a feeling of progression is present. If productive creativity could supplement such feelings of progression, or perhaps *be the progression*, then physical rehabilitation might be a different, more exiting and inspiring part of life after sustained brain injury. Since creativity, challenge/success and motivation are at the core of the original system, the work also focuses on these (all too often ignored) aspects of rehabilitation. In the same perspective this focus might affect the creativity and motivation with which the patient meets every-day tasks. The aim of the project was to create a home based untraditional IT-system supporting these goals and which would work as a supplement to traditional physiotherapeutic training. This is illustrated in figure 1 where the patient is at the top tested in the clinic and the system calibrated with his specific expected progress data. He takes the system home and trains with the family support and without the stress and costs involved of having to report physically to the clinic daily or weekly. The progress of his home training is sent automatically via the internet to the clinic for monitoring, fine tuning and feedback via web cam & mails.

5. CREATIVITY & MOTIVATION

Creativity seems to defy definition and along with motivation is often overlooked in rehabilitation. This means it is often dismissed as a parameter when trying to define measurable goals in rehabilitation (and in many other areas). Yet we know that the feeling of creating something unique or personal is often a source of satisfaction. Thus creativity may well be a source for motivation. In rehabilitation it is clearly apparent, that focusing on improvement promotes motivation for the patient to exercise. With use of feedback to movement of the whole body as well as parts thereof, the project used concepts of music and dance as well as games, tasks and challenges to keep patient's motivation high when 'working out.' A main idea relative to the earlier research by the author is to be able to "hear the way you move" towards an improved proprioception.

6. MOTION CAPTURE

'Motion capture' or 'motion tracking' is about capturing human movement and translating it into knowledge about movement efficiency. It is used for example in sports, film animation and rehabilitation. Several different approaches have been used ranging from expensive multi-tracking camera systems (Vicon, Qualisys, SIMI) to more low cost "wearable systems" (DIEM, Troika) connecting to the computer through cables which are encumbering and impractical for rehabilitation. Such camera systems involve 'expert knowledge and training' for operation and typically are located at institutes which are funded accordingly.

The capture system for the Humanics project was conceptualised as a cost-effective sensor/camera system. Parallel studies with camera systems proved fruitful and it seems pertinent to state that the SoundScapes system that was used for this feasibility study utilising only sensors is now a non-tangible system including cameras and sensors as proposed in the author's original Humanics system design.

7. SOUNDSCAPES SYSTEM

SoundScapes is a system which consists of a variety of non-wearable movement sensors which register motion. The movement information is routed to a computer, which then transforms the data into sound, image and coloured light patterns. Thus, the participants receive an immediate feedback on their movements. In this way, the participants have the opportunity to see/hear their pattern of movement and to see/hear which parts of the body they are moving/not moving. Such a feedback can be of singular importance to for instance people with acquired brain injuries with hemipareses with or without neglect. The participants themselves have the opportunity to choose which kinds of feedback they want.

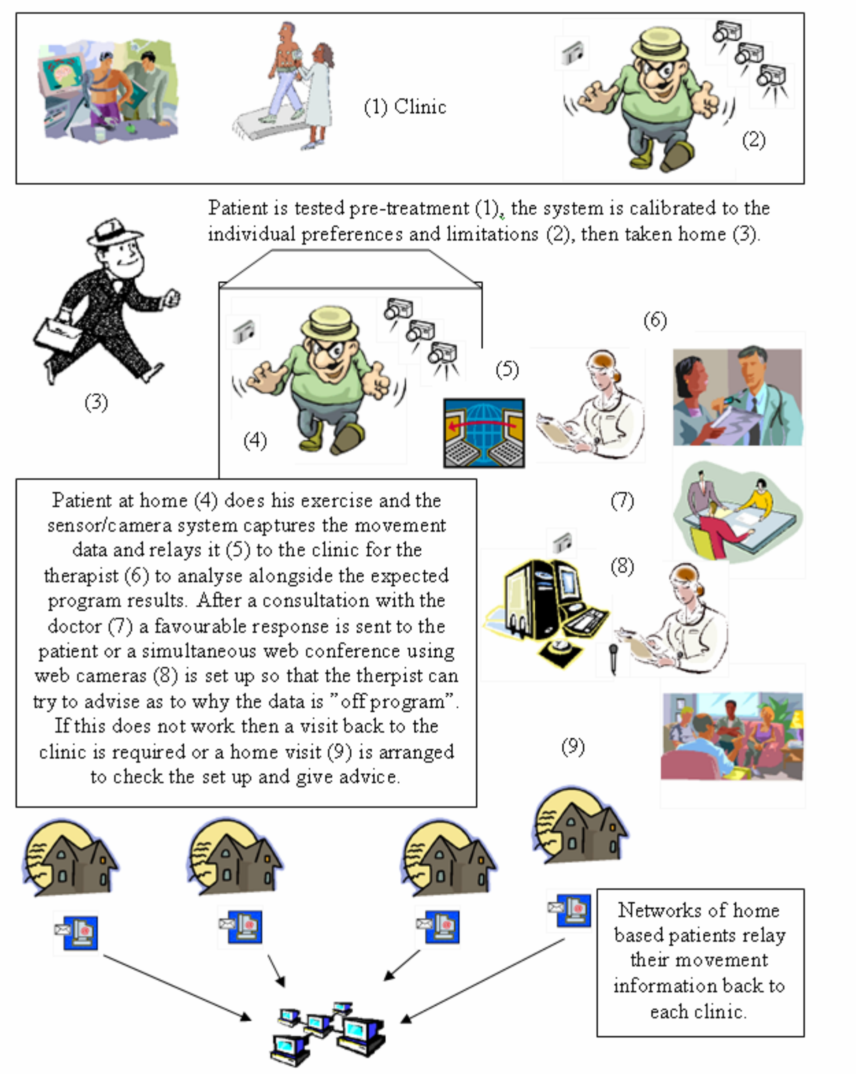


Figure 1. *Humanics telehealth system*

8. RESEARCH OBJECTIVES

The following issues were addressed through the study with the system:

- Does the system give patients with acquired brain injury with physical injuries, an increased physical level of activity and function?
- Does the system have the potential to become a relevant and novel system, and can it increase motivation for physical rehabilitation?
- If the Humanics System is implemented in the patients' home: Is there a potential for increased training efficiency or other benefits by live www-connections between patient and physiotherapist?
- Is implementation of the System as training measurer in private homes a viable prospect?

- When training in the system: Is a more free style of training (e.g. no specific physiotherapeutic exercise or goal) preferable to a more restricted type of training (e.g. specific physiotherapeutic measures and aims, comparing achieved goals, etc.) or vice versa?
- Can results from the system be shown to correlate with functional change measured by traditional physiotherapeutic tests? And if it can; which functions will it be possible/desirable to measure?
- Working with the system: Are creative and motivational aspects of specific activities outside the training sessions affected?

9. METHOD (Phase 1)

The single-case feasibility study included 5 adult patients[i] selected among patients formerly enrolled in the CRBI rehabilitation program[ii]. Training with the baseline SoundScapes system was over one 3-week period. During the training period a rehabilitation physiotherapist collaborated with the author to train and introduce all patients in one-hour sessions. Over the three weeks a maximum attendance was 11. Each session consisted of activities for the patients needs. The entire session of phase 1 utilized the SoundScapes system.

All patients are individually tested following the schedule below:

- At inclusion: User Interface Questionnaire / Prior experience with computers (see appendix 3).
- Immediately before and after the training period. Physiotherapeutic testing including tests of general fitness, balance, level of activity, and quality of movement (see appendix 1). Psychological testing of creativity and motivation (see appendix 2).
- At every training session: Multiple camera video recordings[iii] and audio comments taped by the physiotherapist were collected.
- Where possible results from tests carried out immediately after the traditional rehabilitation program at CRBI are used as reference points.

10. PATIENT SELECTION FOR HUMANICS

As mentioned potential participants were selected from a group of patients formerly enrolled in the traditional post-acute rehabilitation program at the CRBI. At a very general level patients enrolled in this program are typically ½-3 years post-injury and are able to handle most essential ADL at a reasonable level. A total of 51 adults, 26 male and 25 female patients with acquired brain injury (stroke or trauma) aged 24-62 yrs. were selected as potential participants. Exclusion criteria were inherent cerebral dysfunction, any history of psychiatric disease and substance abuse. All participants continue to have physical impairment following their injury (this means that reduced functions of one or more body parts are observed by means of common clinical, neurological assessment). Time post-injury for most patients was 2-7 years having participated in the rehabilitation program 0-4 years ago. Since the patients of the Centre (CRBI) are among the best functioning 30 per cent of all Danish people with brain injuries, many of them were occupied in jobs in some measure. Measures were taken to create homogenous groups of patients in relation to age, gender and localisation / degree of injury. Focus has been put on the largest possible variation of physical after-effects from the injury/the illness.

11. SESSIONS

11.1 Set up

The SoundScapes system was set up as a makeshift laboratory in a large room in the University of Copenhagen adjacent to CRBI. The equipment was primarily - a PC, a three headed infrared volumetric sensor array (author prototype); an ultrasonic linear sensor (*Soundbeam)*; three intelligent robotic light devices; and peripheral interfaces and cabling to source, route and map movement signals into the computer workstation. The protocol of MIDI (Musical Instrument Digital Interface) was the core language.

11.2 Training

An intense three week period was scheduled and the participants attended sessions of one hour duration. Some were not able to attend all of the sessions and the maximum attended possible was 11 times. The most was 10 times with the average of the others coming 5 times over the three week period. The exercises were inductively designed relative to each person's damage and preference for limb/functional training. Many of the improvised exercises proved of great worth and were sustained throughout the period (see 11.3 below.) All were trained with the use of the system giving them an auditive feedback relative to their balance. This was often with the eyes closed which was a problem for many of them and support was required, however they were very positive about this although it was difficult for them. Often they would initiate a sequence of movements that they had previously been instructed to perform or had self-created to help in training. The role of the system in such instances was to give an auditive and/or visual feedback relative to the movement to aid in body awareness. The limbs to be exercised were always located so as to traverse within a Virtual Interactive Space - VIS (Brooks 1999) – volumetric or linear - with a silent 'rest' area adjacent.

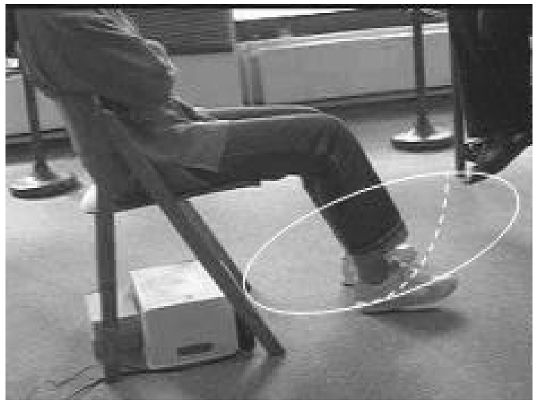


Figure 2: *Extension of leg through interactive sound space exercise: The solid white oval marks the active sense space. The dotted line indicates the foot motion through the active space. The foot starts at the perimeter of the active space and traverses across which results in a scalar tone feedback of a musical instrument, for example a piano. The participant with eyes closed listens to the tone and controls the phrasing and direction of the scale, ascending or descending. This was also programmed to be a familiar melodic tune*.

Figure 3: *Balance exercise: Working with the neglected left side of the patient he faces a large reflector on the wall and closes his eyes as he traverses across the invisible active space towards the author who stands by to assist. The white line marks the body points that active the sounds. This was the best sonic exercise for the male participant as he was incrementally more responsive to visual manipulation as reinforced feedback.*

Figure 4: *Cognitive independence exercise: The movement of the hand in the active (white O) space controls the movement and colour of the light (the triangle on the facing wall) and a 'freeze' pedal under the table out of sight of the participant 'freezes' the position. The participant manipulates to a target on the far wall.*

11.3 Noteworthy incidences

One patient, 63 years of age male who suffered a haemorrhage in the left side of the brain in 1999 had limited observed auditive response in the sessions (figure 3). He informed us that he was a visual artist who had never had any interest in music, and gave a catalogue of his work to the author as a gift in his second session. One visual piece from the catalogue was replicated by the author on the computer so that the man could move his hands between a red, green, and blue filter opening program sensor array and paint the sequence of images that formed the digital version of his piece of art. A major motivation shift occurred through this for him and it is a good example of how the system has to be capable to immediate change to user preferences.

Other exercises that were notably very successful involved the setting up of a sequence (MIDI) of musical tones that constituted a familiar melody and was playable by event triggering through limb movement. This became a phrasing exercise which enabled control relative to desired goal (see figure 2). Another popular exercise was in using an intelligent light scanner that was interfaced to the sensors so that movement of a limb controlled movement of the light image. In this case, with a task given to move the light to a predetermined target usually high on the far left wall, the participant would move a hand in one sensor space to control the horizontal (X) trajectory; then once satisfied to the strategy, they then had to hold that position and press a 'freeze' pedal which was out of sight under the table. Similarly for the vertical (Y) trajectory a movement, a hold and a 'freeze' pedal press. This was perhaps the greatest success as a cognitive independence exercise which was specifically task oriented with a very physical feedback (see figure 4).

11.4 Results

All of those that attended the sessions over the three weeks gave positive response in interviews. They had many ideas following the introduction to the SoundScapes system. Adaptability was a key component so as to be able to fit each individual preferences and limitation. The auditive feedback worked for all but the man who never listened to music mentioned in 11.3 above. All felt that given time they could use the system and thought the home based idea linked by internet a good idea. They also liked the idea of family involvement.

12. CONCLUSION

The feasibility study (phase 1) presented in this paper was a seen by the Danish government body and CRBI as a successful first phase towards the design and realisation of a product and the development contract did in fact follow as a result of this initial study. The small number of participants (5), all positive in the interviews, however could be pointed to as insignificant in number for a research study and the 'loose' methodology implemented in the sessions was not conclusive to a 'hard science' result. The limited time frame of sessions with such a diverse group was also a restriction. A subsequent publication will detail the research for the full product development at CRBI where a larger user group was tested over a longer time period in phase 2.

Acknowledgement: Notably I would like to mention the CRBI ex-Director M. Pinner and physiotherapist G. Thorsen who were the 'Year of The Brain' visitors who initiated the start of the collaboration. Coordinator/secretary L. Lambek, external project coordinator P.K. Larsen, and psychologist E.B. Lyon, all who were with the project for the duration at the Centre - thanks. Others have contributed to the research along the way and whilst I am informed by the CRBI that there is no need to credit names the hard work of Physiotherapist J. Sorensen and Psychologist P. Pipenbring in phase 2 must be mentioned, and similarly H. Mølmark & P. Forster in phase 1. All of the 'students' involved in the sessions and all of the staff at the Centre for Rehabilitation of Brain Injury for giving such fantastic hospitality during the collaborative years when I was based there. Thanks to the Danish '*Erhvervsfremmestyrelsen*' and 'The Egmont Fund,' Denmark who were involved in funding the project, also to my equipment sponsors 'Soundbeam' of Bristol in the UK, IBM for continued support and Martin Lights Denmark for the gratis loan of equipment.

13. REFERENCES AND LINKS

E Asmussen and E Hohwü-Christensen (1980), *Theory of athletics, physiology and kinesiology* (Org. title: Idrætsteori, fysiologi og kinesiologi), Akademisk forlag, Denmark.

A T Beck, C H Ward, M Mendelson, J Mock and J Erbaugh (1961), An inventory for measuring depression. *Archives of Intl. J. General Psychiatry*, 1961, **4**, pp. 561-571.

A T Beck, G K Brown and R A Steer (1996), *Beck Depression Inventory: manual,* The Psychological Corporation, San Antonio, TX.

K Berg, S Wood-Dauphinée, J I Williams and D Gayton (1989a), Measuring balance in the elderly: Preliminary development of an instrument. *Int. J. Physiotherapy Canada*, **41**: pp. 304-311.

K Berg (1989b), Balance and its measure in the elderly: a review. *Int. J. Physiotherapy Canad*a ; **41**: pp. 240

K Berg, B E Maki, J I Williams, S Wood-Dauphinée, and P J Holliday (1992a), Clinical and laboratory measures of postural balance in an elderly population*. Arch. Intl. J. Phys. Med. Rehab*. 1992a; **73**: pp. 1073-1080.

K Berg, S Wood-Dauphinée, J I Williams and B E Maki (1992b), Measuring balance in the elderly: validation of an instrument, *Canadian Journal of Public Health*; **83**: pp.7-11.

K Berg, S Wood-Dauphinée, and J I Williams (1995), The Balance scale: reliability assessment with elderly residents and patients with acute stroke, *Scandinavian Journal of Rehab Med*, **27**, pp. 27-36.

A L Brooks (1999). Virtual Interactive Space, *in Proc. World Confederation for Physical therapy,* Yokohama, Japan; pp. 66.

A L Brooks (2003). *Inhabited Information Spaces 'Living with your data'* – Springer (USA), pp. 89-100.

A W Brown and K Eriksen (2000) *Functional Quality of Movement (FQM) test*, Unp. Copenhagen, Denmark.

I Carlsson, P E Wendt, and J Risberg (2000) On the neurobiology of creativity. Differences in frontal activity between high and low creative subjects. *Intl. J. Neuropsychologia*; **38**, pp. 873-885.

C Cleeland, *Brief Fatigue Inventory*, Pain Research Group, U.T.M.D. Anderson Cancer Centre, Uni. of Texas.

R Crandall (1973) The measurement of self-esteem and related concepts. In J P Robinson and P R Shaver (Eds.), *Measures of social psychological attitudes*, Ann Arbor, MI, University of Michigan Press pp . 45-167.

H J Dupuy (1978) Self-representations of general psychological well-being of American adults, *in proc American Public Health Association Meeting*, Los Angeles, October.

W H Fitts and W L Warren (1996) *Tennessee Self-concept Scale manual* Los Angeles CA, Western Psychological Services.

H Gerhardt (1995) *Rotters incomplete sentences test: Manual*, Dansk Psykologisk Forlag,.

M D Lezak (1995) *Neuropsychological Assessment (3'rd ed.)*, Oxford University Press.

T R Mendoza. X S Wang, C S Cleeland, M Morrissey, B A Johnson, J K Wendt, and S L Huber (1999) The rapid assessment of fatigue severity in cancer patients: use of the Brief Fatigue Inventory. *Cancer,* **85**, pp.1186-96.

M Pinner (2000) Central case management and post acute rehabilitation in Denmark. In A L Christensen and B P Uzzell (Eds.,) *International handbook of neuropsychological rehabilitation.* NY, Kluwer Academic.

R M Reitan and D Wolfson (1985) *The Halstead-Reitan Neuropsychological Test Battery*: *Theory and Clinical Interpretation*. (2nd edn.) Tucson, AZ, Neuropsychology Press.

M Rosenberg (1965) *Society and the adolescent self-image*. New Jersey, Princeton University Press.

J B Rotter and B Villerman (1947) The incomplete sentences test as a method of studying personality, *Journal of Consulting Psychology*, **11**, pp. 43-48.

R C Wylie (1974) *The self-concept: A review of methodology considerations and measuring instruments*, Lincoln, University of Nebraska Press.

APPENDIX 1: PHYSIOTHERAPEUTIC TESTS

Baseline tests: Carried out after the traditional rehabilitation program at the CRBI - *Repeated* before & after test period.

- Grooved Pegboard (Reitan, *et al*., 1992.)
- Halstead's Finger tapping test (Reitan, *et al*., 1992.)
- Åstrands Bicycle ergo meter test – measurement of general fitness (Asmussen, *et al*., 1980.)
- Brief Fatigue Inventory (BFI) (Mendoza, *et al*., 1999. See also BFI hyperlink.)
- Joint flexibility of involved joints (Pain inventory). Manually measured (angle) by physiotherapist.
- Measure of strength of involved muscle groups (Pain inventory).
- Functional Quality of Movement (FQM) inventory in ADL (Brown *et al*., in press).
- Berg Balance Scale - dynamic balance (Berg, *et al*., 1989 a & b; 1992 a & b; 1995).

APPENDIX 2: PSYCHOLOGICAL TESTS AND INTERVIEWS

- Rotter's Sentence completion test (Rotter *et al*.,1947; Gerhardt, 1995.)
- Brick Test[iv]
- The Creative Function Test (Carlsson *et al*., 2000.)
- Tinker Toy Test (Lezak, 1995.)
- Rosenberg Self-esteem Scale, RSES (Rosenberg, 1965; Crandall, 1973; Wylie, 1974.)
- The Tennessee Self-concept Scale, 2'nd ed. (Fitts *et al*., 1996.)
- General Well-Beeing Schedule (Dupuy, 1978, Murrel, 1999.)
- Becks Depression Inventory (Beck, 1961; 1987.)
- Dedicated semi structured interview.
- Focus group interviews including all patients. Led by a psychologist and recorded on video. Experiences, ideas and criticisms are summed up (phase 2 only.)

APPENDIX 3: USER INTERFACE QUESTIONNAIRE

The Questionnaire was designed with open-end questions especially for use with the Humanics project. It addressed two main areas:

- Prior experience with Computers
- User interface of the system (ease of use, feedback, inspiration)

---

[i] One child of 8 years of age recovering from a brain tumour operation also came once and was highly motivated, however sadly due to reaction to chemotherapy he had to stop his visits.

[ii] For further description of the CRBI rehabilitation program and patients see Pinner in Christensen *et al*., 2000

[iii] Evaluation methods pertaining to video recordings - evaluation by 3rd party physiotherapists (minimum of two) are considered the optimal solution.

[iv] This test is often cited as a general measure of creativity. The person being tested is asked within a certain time limit to name as many different purposes for a brick as possible.

NB. The use of (the author owned) equipment in the feasibility study is credited in the report to the Danish authorities as having belonged to the phase 2 commercial contractor Personics. This is incorrect. TB

Interactive virtual environment training for safe street crossing of right hemisphere stroke patients with unilateral spatial neglect

N Katz[1], H Ring[2], Y Naveh[3], R Kizony[4], U Feintuch[5] and P L Weiss[6]

[1,4,5]School of Occupational Therapy, Hadassah-Hebrew University, [2]Loewenstein Rehabilitation Center, and Dept. of Rehabilitation Medicine, Sackler School of Medicine, Tel Aviv University, ISRAEL

[3]Meonot Macabbi-Migdal Hazahav, [4,6]Department of Occupational Therapy, [5]The Caesarea Rothschild Foundation Institute for Interdisciplinary Applications of Computer Science, University of Haifa, ISRAEL

[1]*noomi.katz@huji.ac.il,* [2]*hring@post.tau.ac.il,* [3]*msyuvaln@pluto.huji.ac.il*
[4]*rachelk@zahav.net.il,* [5]*urif@cc.huji.ac.il,* [6]*tamar@research.haifa.ac.il*

ABSTRACT

The goal of this study was to determine whether non immersive interactive virtual environments are an effective medium for training individuals who suffer from Unilateral Spatial Neglect (USN) as a result of a right hemisphere stroke. Participants included 19 patients with stroke in two groups, an experimental group we were given VR-based street crossing training and a control group who were given computer based visual scanning tasks, both for a total of twelve sessions over four weeks. The results achieved by the VR street crossing intervention equalled those achieved by conventional visual scanning tasks. For some measures, the VR intervention even surpassed the scanning tasks in effectiveness. Despite several limitations in this study the present results support further development of the program.

1. INTRODUCTION

The overall goal of this study was to determine whether non immersive interactive virtual environments are an effective medium for training individuals who suffer from Unilateral Spatial Neglect (USN) as a result of a right hemisphere stroke. USN is a major syndrome that affects rehabilitation outcomes (Katz, Hartman-Maeir, Ring & Soroker, 1999).

Although there are various treatment methods that aim to remediate this deficit and to teach compensatory strategies, only some have shown hard evidence of effectiveness. While there is evidence of improvement according to results of impairment assessment, much less evidence exists for increased functional performance at the disability level (Bowen, Lincoln & Dewey, 2003; Cicerone et al., 2000). VR technology may provide a promising method to treat patients who suffer from USN due to its well-known attributes (e.g., ecological validity, the ability to grade the level of difficulty and motivate the user) (Rizzo, & Kim, in press; Rizzo, Schultheis, Kerns, & Mateer, 2004). A PC desk top computer Virtual Reality (VR) technology was the platform of choice in this study as it is more easily available in clinics and rehabilitation units. A screen shot of the initial view of the virtual street is shown in Fig. 1.

Figure 1. *Initial view of street crossing environment.*

The objectives of this study were to 1) to use the virtual street environment developed in a previous study (Naveh, Katz, & Weiss, 2000; Weiss, Naveh & Katz, 2003) to train subjects with USN to become more aware of stimuli in the neglected field of vision and to learn to compensate for their deficit in a safe and graded environment, 2) to compare performance USN measures of subjects with USN who received the VR training to a control group who received computer visual scanning training. 3) to compare subjects' ability to carry out a functional task (crossing a street) prior to and following training in the virtual environment.

2. METHODS

2.1 Participants

Participants included 19 patients in two groups: 1) 11 experimental: 7 men and 4 women, mean age = 62.4 ± 14.0 (SD) years and 2) 8 control: 5 men and 3 women, mean age = 63.3 ± 10.8 years. All participants had a first right hemispheric stroke (RCVA) with persistent USN, 6 to 8 weeks since onset. Included in the study were participants who use any type of mobility aid, but have difficulty in crossing streets in a safe or confident manner. All participants in this study used a wheelchair for mobility.

2.2 Instruments

2.2.1 Virtual environment. A street crossing virtual environment was programmed via Supers cape's 3D Webmaster and run on a desktop computer, with successively graded levels of difficulty that provide users with an opportunity to decide when it is safe to cross a virtual street. It was initially tested on 12 subjects, six patients with stroke and six matched controls (Naveh et al., 2000; Weiss et al., 2003). Results showed that the program is suitable for patients with neurological deficits in both its cognitive and motor demands. In the current study the virtual street was used to assess the performance within the environment as well as a treatment tool.

2.2.2 USN measures. Severity of USN was measured with conventional paper and pencil measures, Star cancellation from the Behavioral Inattention Test (BIT) (Wilson, Cockburn & Halligan, 1987), Mesulam Symbol Cancellation test (Mesulam, 2000). For each of these tests the number of items cancelled correctly on the left side was calculated and the time it took to complete the test was measured. In addition an ADL checklist was used to measure the affect of the USN on daily activities. For this measure lower scores represent higher performance (Hartman-Maeir & Katz, 1995).

2.2.3 Functional performance. Functional evaluation of the ability to cross an actual street was evaluated from video-taped records. Since the participants were in wheelchairs, an occupational therapist wheeled their chair to cross the road only when they told her that it was safe to do so. The street crossing was performed on a divided road so that the participants first had to cross the lanes where vehicles were approaching from left to reach a central island. They then continued to the other side where cars came from right. Participants then re-crossed the road to return to the start point. In this way we observed and videotaped two pre- and post intervention street crossings, with vehicles approaching the crosswalk twice from the left and twice from the right. Since an essential difficulty for right hemisphere stroke patients with USN is the leftward gaze and consideration of objects on the left, their ability to look to the left was considered to be an especially critical indicator of safety during street crossing.

Since at any given point in time the number of vehicles passing on a road varied, performance scores were normalized by dividing the time it took for a patient to decide to cross the road with the number of vehicles that passed during that time. This measure was termed the "Decision time per vehicle prior to initiation of street crossing". An additional score was the number of times the participant looked to the left.

2.3 Procedure

The study took place originally at Loewenstein Rehabilitation Hospital, with additional participants from Hadassah Medical Center rehabilitation department and Beit Rivka Medical and Geriatric Center. The study was approved by the Institutional Review Boards of the respective hospitals. In each facility participants were randomly assigned to either group: VR training (experimental) and computer visual scanning tasks (control).

In the current controlled clinical trial, the virtual reality training protocol continued for 4 weeks, with 3 sessions per week, each of 45 minutes duration, for a total of 9 hours. The timing of the control group computer scanning training protocol was identical. Prior to commencement of training and subsequent to it, all participants were assessed with the USN measures, a virtual street crossing test and a real street crossing test.

2.4 *Data analysis*

The data were analyzed using group means to compare pre-post differences within each group and between groups at each testing time. In addition, in order to demonstrate the change occurring due to intervention, we calculated and compared the mean pre- and post differences for each variable. A one tailed t-test was used as the VR experimental group was expected to show greater improvement than the control group.

3. RESULTS

As is the case for most desktop VR systems, no cybersickness-type side effects were noted for any of the participants; all enjoyed using the program and willingly participated in the intervention. The pre- and post-test results for both patient groups are shown in the Table and include scores for the USN measurements, the VR measurements, and the real street crossing measurements.

3.1 *USN measures*

We first noted that the VR group performed lower at pre-test on most measures which is indicative of a more severe USN. Thus, although participants were randomly assigned to each group, the USN for the VR group was, on average, more severe. Both groups, regardless of intervention type, improved in their scores on the paper and pencil tasks, namely the number of correctly cancelled items on the Star cancellation and the Mesulam random cancellation test. The difference for the Mesulam was significant at $p<.01$ as indicated in the Table by the two interconnected arrows.

Both groups took less time to complete the test following intervention but the difference was not significant. It is interesting to note that, on average, the VR group required much more time to complete the Mesulam during the pre-test (mean = 519 s) than did the control group (mean = 353.9 s). However, the mean pre-post test difference for the VR group is more than twice as large as that of the control group (137.5 s versus 52.7 s). Although there were no pre-post significant differences in the Star cancellation scores (number of stimuli correctly cancelled), the VR group took less time to complete this task at post test whereas the control group needed longer time to complete it at post-test. Performance on the ADL checklist, that is a patient's ability to cope with daily living skills (which reflect the functional implications of USN) showed significant improvement for both groups from pre- to post-test ($p<.05$). Note that lower scores indicate improvement on this measure. This result together with the cancellation tests shows that patients from both groups decreased in the severity of USN, although all patients still showed persistent USN at time of the post-test.

3.2 *VR Street crossing measures*

The VR street crossing performance of both groups showed the effects of training. Specifically, there was improvement in the number of times participants looked to the left for the VR group; this difference was significant at $p<.05$. More importantly, in the VR group, most patients made fewer accidents (about 50%) during the virtual street crossing at post-test which was significant at $p<.035$ (one tailed). In contrast, only one member of the control group had fewer accidents while the others did not change their performance from pre- to post-test. Their mean pre- and post-test number of accidents was similar. Comparing the two mean differences (4.1 to -0.2) between groups was significant at $p<.035$ (one tailed, see Table).

3.3 *Real Street crossing measurements*

Pre- and post-test real street crossings were videotaped and then analyzed using the two measures indicated above, the number of times a person looked to the left and the decision time per vehicle prior to initiation of street crossing. For the VR group, the mean number of times participants looked to the left before crossing increased from pre- to post-test, whereas for the control group this number decreased slightly. This difference was not significant but the direction of the results indicates that there was greater improvement for the VR group. The difference between the two groups is highlighted in Fig. 2 which shows that a greater number of participants in the VR group looked to the left at post-test as compared to the control group who demonstrated no change. The decision time to cross the street per vehicle showed no change in the VR group and a slight decrease in the control group. Mean differences were not significant.

Table 1. *Measures of USN, VR street crossing test and real street crossing. Note that arrows between two means indicate significant results.*

	VR Group (N=11)			Control Group (N=8)		
	Pre-test Mean(SD)	Post-test Mean(SD)	Difference Mean(SD)	Pre-test Mean(SD)	Post-test Mean(SD)	Difference Mean(SD)
Attention measures						
Star Cancellation						
Score – # correct Left	9.2 (9.7)	14.8 (12.9)	-5.6 (11.0)	14.6 (10.4)	18.1 (10.2)	-3.5 (4.8)
Time in seconds	249.8 (211.8)	213.3 (137.6)	36.5 (149.1)	181.1 (112.7)	194.2 (57.5)	4.8 (124.5)
Mesulam Cancellation						
Score – # correct Left	7.4 (9.2)	13.6 (12.2)	-6.3 (6.4)	6.5 (9.3)	12.6 (10.6)	-6.1 (6.4)
Time in seconds	519.0 (559.7)	381.5 (326.3)	137.5 (430.0)	353.9 (199.6)	290.3 (61.9)	52.7 (224.9)
ADL Checklist						
Therapist score	2.2 (0.5)	1.4 (0.6)	0.8 (0.3)	1.4 (0.7)	0.8 (0.5)	0.6 (0.4)
VR measures						
# Look left	10.5 (5.0)	17.3 (7.2)	-6.2 (5.6)	7.8 (6.5)	14.4 (13.9)	-5.2 (10.3)
# of Accidents	7.9 (6.9)	3.8 (4.6)	4.1 (6.8)	3.8 (2.8)	3.4 (2.7)	-0.2 (1.8)
Real Street Crossing measures						
# of times look left	4.0 (2.4)	5.4 (2.9)	-1.4 (2.8)	6.3 (3.7)	5.8 (4.6)	0.5 (6.2)
Decision time per vehicle prior to initiation of street crossing						
when vehicles come from left	5.6 (7.9)	5.7 (7.6)	-0.2 (9.0)	12.8 (14.1)	10.1 (9.2)	2.7 (6.3)
when vehicles come from right	9.6(12.1)	7.3 (7.7)	2.2(13.7)	8.2(3.9)	6.1(5.6)	2.0(6.8)

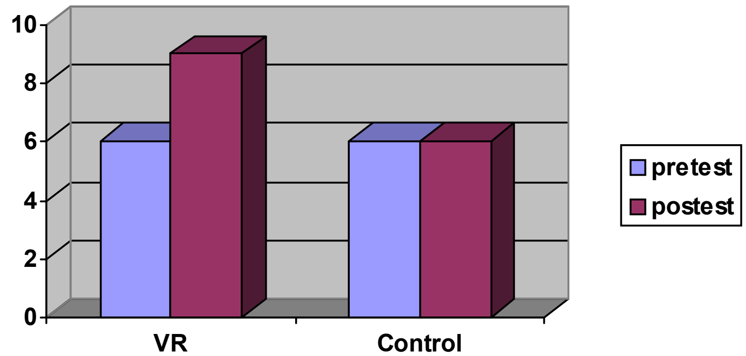


Figure 2. *Histogram showing frequencies of how many patients in each group looked to the left in the real street crossing test.*

4. DISCUSSION

The results of the study show that the VR intervention was effective both in terms of improving visual-spatial performance as measured in this study and for some improvement in the ability to cross a real street. It therefore appears to have potential to become a useful tool in rehabilitation. Although the results of some of the analyses did not reach significance at the 0.05 level (likely due to the small sample size) the direction of improvement was evident. In addition, the fact of the initial pre-test difference on the severity of USN between the groups which occurred despite the random assignment to intervention type was a limitation. We attempted to correct for this difference by comparing pre- and post-test differences rather than absolute scores. Nevertheless, these differences make it more difficult to assess the effect of VR intervention as compared to conventional computer scanning tasks.

It is important to note that treatment based on the computer visual scanning tasks that were provided to the control group in this study are, according to most research evidence for USN therapy, currently the intervention method of choice. As a result of a meta analysis reviewing the effectiveness in the treatment of visual spatial deficits, Cicerone et al. (2000) recommended that training visual scanning be used as a 'practice standard' to improve patients' ability to compensate for visual neglect after right hemisphere stroke. In addition, the evidence demonstrated that training on more complex tasks appears to enhance patients' performance and facilitate generalization to other functional areas as well. Similar recommendations were made by Cappa et al. (2003) in their report.

The results achieved by this VR street crossing intervention equalled, at the very least, those achieved by conventional visual scanning tasks. For some measures, the VR intervention even surpassed the scanning tasks in effectiveness. This is an important finding given the fact that this virtual environment, as well as the simple desktop hardware used to display it, constitutes the lower end of the immersion spectrum. Until recently, there was only a limited capacity to grade the difficulty of the street crossing task and to control the delivery of stimuli in accordance with a given patient's ability. Taking these limitations into account it seems that the use of this VR street environment is encouraging and supports further development of the program.

The above considerations, together with observations of the video-taped records, and discussions with the clinicians involved in the VR training process have led us to revise our environment and experimental paradigm. We are currently carrying out a second phase of the study and are including patients who are at a more advanced stage of rehabilitation. That is, post-acute patients who participate in rehabilitation as out-patients and have the need to cope with real-life street crossing as a routine daily task. This population would likely be even more motivated to use this tool, and stand to obtain greater benefit from it. Thus the second phase of the study includes participants with the same diagnosis at ambulatory and post-hospitalization or day treatment. Participants are able to walk alone or with assistive devices such as a cane, tripod or walker and will thus be able to be tested during independent street crossing.

Moreover, we have added a questionnaire of pedestrian walking behaviours, both prior to the stroke and since the stroke. This will enable us to determine how improvements in USN, as well as virtual and real street crossing skills are correlated with the participants' functional abilities. Finally, the original street crossing program has been revised to include a greater variety of pedestrian situations (e.g., traffic lights), different routes to choose from and places to go to (see Fig. 3). The new version provides opportunities to explore the participants' executive functions, in particular as related to planning and decision making. In addition to the current population of older adult patients with stroke, this new environment is currently being tested to train pedestrian safety for children with autism.

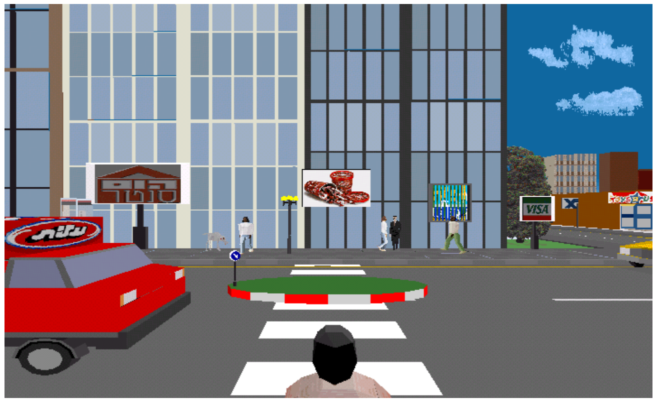


Figure 3. *Screen shot showing revised street crossing environment.*

5. REFERENCES

A Bowen, N B Lincoln, and M Dewey (2003) Cognitive rehabilitation for spatial neglect following stroke (Cochrane review). In: *The Cochrane Library*, Issue 4. Chichester, UK: John Wiley & Sons, Ltd.

S F Cappa, T Benke, S Clarke, B Rossi, B Stemmer and C M Van-Heugten (2003) EFNS guidelines on cognitive rehabilitation: report of an EFNS task force. *Europ J Neurol, 10,* pp. 11-23.

K D Cicerone KD, C Dahlberg, K Kalmar, D M Langenbahn, J F Malec, T F Bergquist, T Felicetti, J T Giacino, J P Harley, D E Harrington, J Herzog, S Kneipp, L Laatsch and P A Morse. (2000) Evidence-Based Cognitive Rehabilitation: Recommendations for Clinical Practice. *Arch Phys Med Rehab, 81*, pp. 1596-1615.

A Hartman-Maeir and N Katz (1995) Validity of the Behavioral Inattention Test (BIT): Relationships with task performance. *Am J Occupat Ther, 49*, pp. 507-511.

N Katz, A Hartman Maeir, H Ring, and N Soroker (1999) Functional disability and rehabilitation outcome in right-hemispheric-damaged patients with and without unilateral spatial neglect. *Arch Phys Med Rehab, 80,* pp. 379-384.

N Katz, H Ring, Y Naveh, R Kizony, U Feintuch and P L Weiss (2003) Effect of interactive virtual environment training on independent safe street crossing of stroke patients with unilateral spatial neglect. Presentation at the Israeli Association of Physical medicine and rehabilitation conference. Shfaim, Israel.

M Mesulam (2000) *Behavioral Neuroanatomy. Principles of behavioral neurology*. Philadelphia: F.A. Davis.

Y Naveh, N Katz and P L Weiss (2000) The effects of interactive virtual environment training on independent Safe Street crossing of right CVA patients with unilateral spatial neglect. Proceedings of The *3rd International Conference on Disability, Virtual Reality and Associated Technology*. Italy: Alghero. pp. 243-248.

A A Rizzo, M T Schultheis, K Kerns and C Mateer (2004) Analysis of assets for virtual reality in neuropsychology. *Neuropsych rehab, 14*, pp. 207-239.

A A Rizzo and G J Kim (in press) A SWOT Analysis of the Field of VR Rehabilitation and Therapy. *Presence: Teleoperators and Virtual Environments*.

P L Weiss, Y Naveh and N Katz (2003) Design and testing of a virtual environment to train CVA patients with unilateral spatial neglect to cross a street safely. *Occupat Ther Int, 10*, pp. 39-55.

B A Wilson, J Cockburn and P W Halligan (1987) *Behavioural Inattention Test Manual*. Fareham, Hants, England: Thames Valley Test Co.

Using virtual public transport for treating phobias

C Sik Lanyi[1], V Simon[2], L Simon[3] and V Laky[4]

[1,4]Department of Image Processing and Neurocomputing, University of Veszprem,
Egyetem u. 10., H-8200 Veszprém, HUNGARY

[2,3]Department of Psychiatry and Psychotherapy, Semmelweis Medical University,
Balassa u. 6. H-1083 Budapest, HUNGARY

lanyi@almos.vein.hu, *simonviktoria@yahoo.com, simon@psych.sote.hu, Viktoria.Laky@hu.michelin.com*,

[1,4]*www.knt.vein.hu,* [2,3]*www.sote.hu*

ABSTRACT

Nowadays using and talking about virtual reality (VR) is a very popular subject. VR is an artificial world, a computer mediated environment. The user tries to enter fully into the spirit of her or his role in this unreal-world. Virtual Environment (VE) technology has undergone a transition in the past few years that has taken it out of the realm of expensive toy and into that of functional technology. During the past decade, in the field of Mental Healthcare, the considerable potential of VEs has been recognised for the scientific study. This paper shows the application of VR and presents the VR research in the University of Veszprem. The virtual worlds, introduced below, are developed for treating specific phobias (fear of travelling).

1. INTRODUCTION

The medical application and research of VR has more than one and a half decades of history. For all this the cooperation of more scientific fields are needed. Beside computer science and software engineering - that provides technology - also sociology, psychology are needed especially to reveal the psychological background of the effectiveness of VR and either to improve it.

VR is a computer mediated environment where one can step in, move and also create an interactive connection with. The psychological background of its effectiveness as mentioned above is not yet clear in details. What we surely know by now that one of the keywords of its effectiveness is “presence”. Presence is feeling “real” in the mediated world. It is a product of the mind, created by a complex psychological procedure. This psychological procedure is closely related to self-perception and attention preferences also play very important role.

The number of fields for using VR are constantly expanding. From the first applications in games, entertainment through army applications nowadays healthcare and education turned to be in the middle of interest. They use VR for the training of the aircraft, pilots, firemen, divers, soldiers etc. In medicine, VR starts to play role in the training of surgeons, medical students and in the field of psychiatry, psychotherapy, neuropsychiatry. The latest applications are pain reduction and physiotherapy, rehabilitation and scientific researches.

In neuropsychiatry learning memory dysfunctions, learning and improving cognitive dysfunctions, helping Parkinson-patients with movement-induction, the rehabilitation of post-stroke patients are the most improving fields. In children’s psychiatry researches have been shown in the field of improving children with autism, ADHD and other attention disorders. In psychotherapy VR is already proved to be effective in the therapy of anxiety and body-perception problems, PTSD and phobias.

Next to clinical experiences many controlled clinical researches have already proved the equal effect of desensitisation with VR and regular psychotherapy. Wiederhold and colleagues have published the analysis of a long term, prospective research with the biggest sample yet. It e.g. shows the already mentioned equality between the two kinds of therapies but also shows that those therapies where next VR bio-feedback was also used were even more effective and what is really interesting and important that only 4 % was the dropout during the 4 Years of data collecting.

Advantages of treating phobias with VR are:

- it is easier to complete the tasks in the office instead of organising complicated outside programs
- it is safer and more controllable by the therapist who can monitor the effects and react to them more precisely.
- It is repeatable
- Easier to engage the patients for the therapy and the dropout is less during research or therapy

Disadvantages of the therapy with VR include

- simulator sickness may occur
- still expensive

2. THE PROJECT

Phobia is an anxiety disorder in which the person has an extreme, unrealistic fear of a particular situation, activity or object. The fear over which has no control occurs whenever the object appears or the situation arises and the phobic man experiences strong, unbearable anxiety and distressing symptoms such as heart palpitations, sweating or feeling of panic when presented with the feared object or situation. (Beck, 1976; Laky, 2003; Szendi, 2000)

The person with phobia makes real efforts in order to avoid the feared situation or object. Based on the feared object or situation and the strength of the avoidance phobia can make one's life unbearable, others can live with their phobia without serious everyday-problems. For example in Hungary if somebody has snake-phobia is not a real problem since unless you go to the zoo it is hard to meet snakes. (Of course anytime a neighbour can move in with his beloved pet..) What is important that in this case avoiding the feared animal doesn't need extra energy and suffering. In another case for example with travel phobia somebody in Budapest can reach a point when he is not even able to leave his flat.

Fear of public transport is a part of the agoraphobia and its object could be travelling by train, bus, tram, aeroplane, car and the most frequented (at least in Hungary) is fear of travelling by underground.

The method of treating travel phobia with VR is practically the same as the regular psychotherapy, called desensitisation. Before the exposition with the feared situation patients learn relaxation and other methods for eliminating anxiety and they also have to create a hierarchy of the stimuli in order to show which causes the less and the strongest anxiety. After this the exposition starts and with VR the therapist can control, based on the hierarchy, the number and the quality of stimuli that the patient has to cope with.

This paper shows a pilot VE intended for people who fear of travelling by underground.

3. DEVELOPING THE VIRTUAL ENVIRONMENT

The model of the developed virtual environment is one of the Hungarian underground located in Budapest. First we prepared videos about this underground and after according to these videos we created the virtual underground by the help of Maya software. The following pictures show the real and the modelled escalator of the underground (Figures 1 and 2) and the wire frame model of the station (Figure 3).

After modelling the underground by Maya we exported the virtual objects to Shockwave 3D file format (the extension of this file is W3D). We had to do it for getting the VE interactive. For the exporting we use Shockwave Exporter that could be found free on the Web. This program is able to export Maya made object to the format (W3D) could be easily used by Multimedia Director.

Multimedia Director helps us to make our underground interactive. For this aim we wrote a small program which can modify the viewpoint of the camera by rotating 3 dimensional vectors. It permits the user to look around without limitation. The walk about is not allowed, because for it we should have had to do impact monitoring that demands a lot of computing result slowing down the computer(s). So the user while "walking" along a given route, by the help of predefined camera moving, is able to make a stop whenever he likes and he can look around. (The limitation of free walking is not a hard restriction since the escalator is limits the moving in itself.)

Figure 1. *The real escalator.*

Figure 2. *The virtual escalator.*

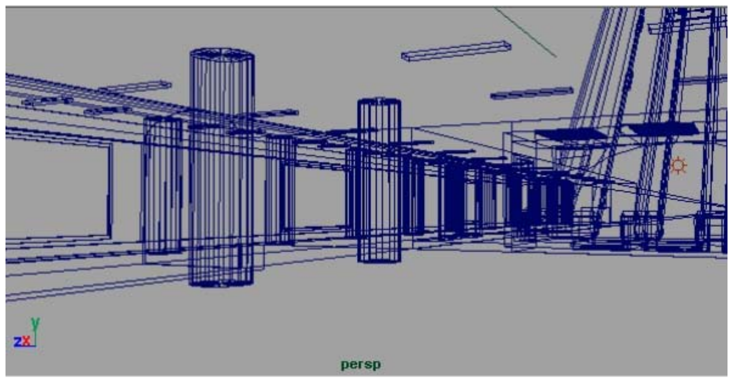


Figure 3. *The wire frame model of the station.*

For the VE we clipped the sounds from the real video for elevating the feel of reality. These sounds were applied for the prepared virtual world. The voices bulk large that the patients feel like being on the escalator or getting on a block. Some special sound effect, for example the voice indicates the departure of the underground, have particular part in that sufferer feels the same anxiety as he feels during travelling on real ones. The importance of sounds is so considerable that it could decrease the magnitude of photo-realistic graphics.

4. THE PROGRAM

The following pictures show some picture about the result, the escalator (Figure 4), the station (Figure 5) and the underground carriage outside (Figure 6) and inside (Figure 7).

The usage of the program is not so hard. The mouse controls the viewpoint of the user and there are some function key and/or button for exiting, stopping, etc. It takes only few minutes that the user makes himself at home in the program. The user has to choose between window and full screen mode and the therapy starts. We put a lot of stress on the program being easy to use. It is important that both the patient and the therapist can use the program easily so that they can turn their full attention to the task and the therapeutic situation. During planning the software we concentrated especially on making it possible to be easy to follow the most important rule of the desensitisation process of behavioural therapy, the rule of “step by step”. The flexibility of the program gives wide possibilities of the length, strength and time of the exercise during the therapy. It is even possible to step one level back if the actual task causes too high level of anxiety – even during relaxation- to the patient. Practically the exercise can be repeated on a certain level without restrictions. The patient or the therapist can stop the program whenever they feel like it. After stopping the program it can be restarted or continue.

During preparing the software we didn’t put stress on its image being full of with details since without the details VE is a proper surface for the patients to project their former experiences and anxiety memories.

Requirements: Pentium class or equivalent CPU (2 GHz), SVGA display with 800*600 resolution, quick video card (minimum Ati Radeon 9000) Windows XP operating system Mouse, Sound card.

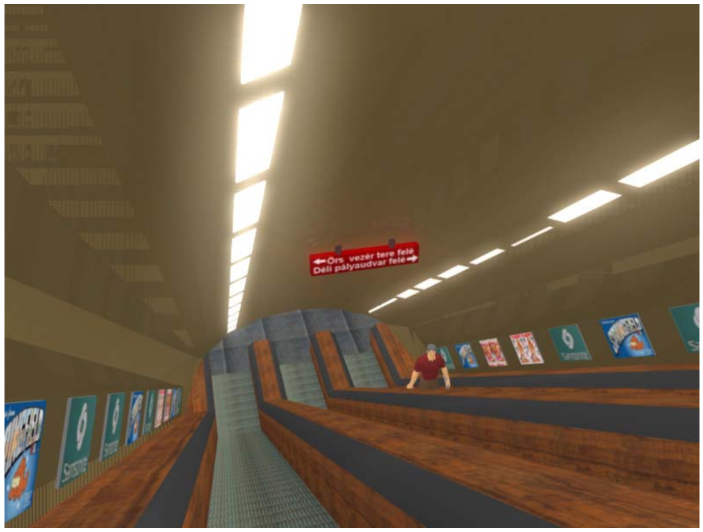


Figure 4. *The escalator*

Figure 5. *The station*

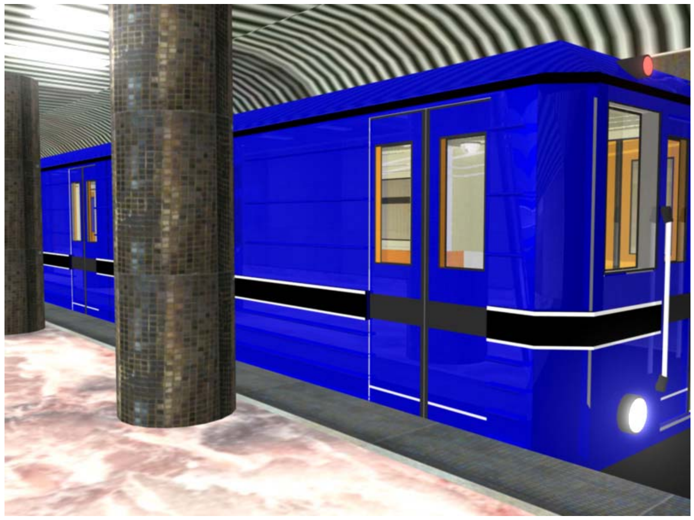


Figure 6. *The underground carriage outside*

Figure 7. *The underground carriage inside*

5. TESTING

The tests of the applicability of this program in the therapy of patients with phobia will be placed at the Department of Psychiatry and Psychotherapy in Semmelweis Medical University. Our first experiences with patients and healthy people shown the software is that they use the program without any difficulty and report about very reality-like feeling, though they miss the possibility to move freely in the environment. These tests were not clinically controlled examinations only helping us to have the first package of information about practical features and usability of the software. We plan a clinically controlled study with two steps. During the first step we are going to measure the feeling of presence and state anxiety when using the VE among phobic patients and matched control subjects. Our goal is to show that entering the VE the patients experience the same anxiety level as when they are exposed to the feared stimuli in vivo so it creates a proper basis for psychotherapy with desensitisation. During the second step we would like to start therapy with VR on a small group of pilot cases in order to show the – at least – equal effects of desensitisation with VR and regular psychotherapy and also to show the upper mentioned practical advantages of psychotherapy in VE.

6. CONCLUSIONS & FUTURE PLANS

1. Our paper introduced a new VE created for treating travel phobia
2. VR has not only useful entertaining and educational applications but proved to be important in the field of medicine especially in psychiatry. It can be a good tool to strengthen the effect of cognitive and behavioural psychotherapies
3. It is important to improve this application of VR both from the aspect of computer science and psychiatry

Our aims to the future is modelling and creating new VEs for treating other public transport phobias, such as fear of travelling by tram or bus.

Acknowledgements: The authors would like to acknowledge the Alias Wavefront to ensure free use of Maya software for 5 PCs to University of Veszprem. We would also like to acknowledge the help of university student Mr Norbert Posztos and Mr Gabor Toth, who carried out the real video recording and Maya modelling.

7. REFERENCES

A Beck (1976), *Cognitive Therapy and Emotional Disorder,.* Penguin

M E Domingo, J C G Cebollada, C P Salvador (2001), VR Testbed Configuration for Phobia Treatment Research, *Proceeding of the Euromedia'2001 Conference,* pp. 200-204.

H R Lückert (1997), *Félelem és pánik*, Trivium Kiadó

L F Hodges, P Anderson, G C Burdea, H G Hoffman, B O Routhbaum (2001), Treating Psychological and Physical Disorders with VR, *IEEE Computer Graphics and Applications*, November/December, pp. 25-33.

G Kabdebó (1995), A látszólagos valóság, *A Természet Világa* 125. 9., pp. 401-404.

V Laky (2002), Virtual Reality in the Psychology. *MSC Thesis Work University of Veszprem,* Veszprem

V Laky, C Sik-Lányi (2003), Using virtual reality in psychology (Virtual worlds in treating agoraphobia and acrophobia), *7th European Conference for the Advancement of Assertive Technology in Europe*, Dublin, pp. 628-632.

V Laky, C Sik-Lányi (2003), To develop Virtual Reality worlds for treating agoraphobia and acrophobia. *VRIC Virtual Reality International Conference,* Laval, pp. 127-133.

M M North, S M North, J R Coble (1998), Virtual Reality Therapy: An Effective Treatment for Psychological Disorders, *Virtual Reality Neuro-Psycho-Physiology*, IOS Press

M Pesce (1995), *VRML: browsing and building cyberspace*, New Riders Publishing, Indianapolis

F N Rahayu (2003), *Virtual Reality for Social Phobia and Agoraphobia Treatment,* Informatietechnologie en Systemen Technische Universiteit Delft

G Riva (1998), *Virtual Reality as Assessment Tool in Psychology,* Virtual Reality Neuro-Psycho-Physiology, IOS Press

G Riva (2001), *Communications Trough Virtual Technology: Identity Community and Technology in the Internet Age*, IOS Press

L Rosenblum, M Macedonia (2001), Adorning VRML Worlds with Environmental Aspects, *IEEE Computer Graphics and Applications,* January/February, pp. 6-11

M J Schuemie, Design of Virtual Reality Exposure Therapy System – Task Analysis

M J Schuemie, C A P G van der Mast, Presence: Interacting in VR?

M J Schuemie, C A P G van der Mast, M Bruynzeel, L Dorst, M Brinckman, G Emmelkamp, P de Haan, Treatment of Acrophobia in Virtual Reality: a Pilot Study

G Szendi (2000), Pánik betegség: a probléma közel sincs megoldva. *Praxis* 9 (10), pp. 13-17.

J Vacca (1996). Virtual Reality. Computer Technology Resource Corp.

Preliminary evaluation of a virtual reality-based driving assessment test

F D Rose[1], B M Brooks[2] and A G Leadbetter[3]

School of Psychology, University of East London,
Romford Road, Stratford, London, ENGLAND

f.d.rose@uel.ac.uk, b.m.brooks@uel.ac.uk, a.g.leadbetter@uel.ac.uk

[1,2,3]*www.uel.ac.uk*

ABSTRACT

Assessing one's own driving ability is very subjective, and there are occasions when an objective off-road assessment would be very useful, and potentially life-saving. For example, after physical or mental trauma, or approaching old age, it would be very useful for people to perform their own off-road assessment to help them to decide whether they should resume driving, or continue to drive. It is possible that people might be more likely to accept that it would be inadvisable for them to drive if they had themselves performed such an assessment. We are currently evaluating a virtual reality (VR) based driving assessment which runs on a PC and could be made easily accessible to people in these circumstances. The first stage of the evaluation was to evaluate the performance of drivers and non-drivers on the VR driving assessment and to compare the results obtained across the two groups of participants and with their performance on the Stroke Drivers Screening Assessment (SDSA). The VR driving assessment discriminated between drivers and non-drivers but the SDSA did not. In addition, two measures on the VR driving assessment correlated with drivers' scores on the SDSA.

1. INTRODUCTION

In recent years, Virtual Reality (VR) has been used to good effect in numerous training situations. Among the advantages of virtual training is that it allows individuals to work at their own speed and in situations tailored to their own specific needs. Less emphasis has been placed on the value of virtual environments in allowing people to arrive at realistic assessments of their own abilities with regard to the particular task. Sometimes, such a self assessment may inform the type of further training needed. On other occasions, it may be just as valuable in persuading the person against continuing with the activity in question. The VR application described here is primarily concerned with helping people to assess their own performance levels.

Older people often see the continuation of their ability to drive as a crucial factor in maintaining their independence (Barnes and Hoyle, 1995; Mollenkopf et al., 1997, Rabbitt et al., 1996). Similarly, drivers who have suffered physical or mental trauma often consider that being able to drive again is an important landmark in their recovery and return to normal life (Mazaux et al., 1997). Currently, however, decisions as to whether older people should continue to drive, or whether people should return to driving after physical or mental trauma, are taken on an ad hoc basis, despite potentially disastrous consequences if the wrong decision is made (British Psychological Society, 1999). Regrettably, there is no standardised method whereby older drivers or drivers who have suffered trauma can assess their own ability to continue or resume driving. VR has the potential to provide the means whereby people can test their own driving ability without exposing themselves to the potential dangers of real driving. Many skills associated with real driving can be assessed on a VR-based driving assessment test which runs on a PC. On completion of the assessment, participants could be given a print-out of their performance and a recommendation as to whether or not they should continue or resume driving. This type of driving assessment would be economical and easy to install in driving assessment centres, public libraries, Age Concern centres and rehabilitation wards, for example. Although such an assessment would not replace any safeguards that are currently in place, it would allow people more informed control over their own lives and would have the potential to save lives by making some people aware that their standard of driving is dangerous.

We have devised a VR-based driving assessment test that runs on a PC. A screen-shot from the virtual environment is shown below at Fig. 1. In the assessment, a virtual car is driven using a widely-available steering wheel, accelerator and brake set. The test requires participants to follow a sign-posted route along

ISBN 07 049 11 44 2

major and minor roads to a particular destination (a zoo). Along the route are road junctions, traffic lights and roundabouts which the participant is required to negotiate whilst complying with road signs and speed restrictions, and interacting with other traffic.

Figure 1. *A screen-shot from the VR driving assessment test.*

We have recently begun an evaluation of our VR driving test which will be in three stages. In the first stage, the performance of drivers and non-drivers will be compared on a prototype of the test. The justification for this evaluation stage is that the VR driving test should be capable of measuring some of the driving skills that are acquired with experience. It should therefore be capable of differentiating between the performance of drivers and non-drivers. Performance measures that do not produce a difference between drivers and non-drivers will be improved and incorporated into a revised version of the test. In the second stage, the test will be performed by drivers who have suffered a stroke and older drivers to investigate whether they are able to understand the test and can use it independently. Again, any alterations required will be incorporated into a revised version. In the third stage, volunteer older drivers' performance on the test will be compared with their performance on a real DSA driving test conducted by an approved driving instructor. The VR driving assessment test will continue to be revised and evaluated until there is a high correlation between participants' performance on the test and their performance on a real world driving test. The first stage of the evaluation is described below.

2. METHOD

1.1 Participants

Forty students from the University of East London voluntarily participated in the study to earn course credits. Twenty of the students had a full driving licence and 20 had never driven a car. The two groups of participants did not differ significantly in age [Drivers - age range 18-52, mean age 28 years; Non-drivers - age range 18-54, mean age 25 years, t(38)=0.85, p=0.40]. There were 18 female and 2 male drivers and 17 female and 3 male non-drivers. Ten drivers and 7 non-drivers had previously used a driving simulator.

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

1.2 Equipment and Materials

The VR-based driving assessment test was constructed using World-Up software and run on a Duel Xeon PC with a 17” monitor using a *Logitech* steering wheel, accelerator and brake set. Buttons on either side of the steering wheel allow the user to look left or right. The assessment comprises a sign-posted route to the zoo with left-hand and right-hand turns, two sets of traffic lights and two roundabouts. The majority of the traffic in the simulator runs pseudo-randomly but some traffic is programmed to produce potentially hazardous traffic situations which the participant has to negotiate. A record is kept of participants’ performance, including how often they crash and hit the kerb, their steering ability, their stopping times, how long they take to merge with other traffic, how well they conform to the speed limit, and the total time they take to complete the route (approximately 10 minutes, depending on the user’s ability). A practise route comprising a circular road with one set of traffic lights enables users to familiarise themselves with steering the virtual car and operating the controls. A photograph of a participant using the VR driving assessment is shown at Fig. 2.

Figure 2. *A participant using the VR driving assessment test.*

The participants also performed the Stroke Driver Screening Assessment (SDSA) (Nouri & Lincoln, 1994). The SDSA comprises four cognitive tests which have been compiled to evaluate driving fitness in stroke patients and have been shown to be significantly better than other forms of driving assessment in determining individuals who were found to be unsafe to drive (Nouri & Lincoln, 1993). Three of the four tests are used in the overall assessment. The first is a measure of attention and concentration in which participants are presented with a sheet of dots in groups of three, four or five. The task is to cross out each group of four dots within a maximum time of 15 minutes. The second test is a measure of non-verbal reasoning in which participants are required to correctly place 16 cards depicting two vehicles travelling in different directions from a roundabout into a grid of compass directions so that the direction of each vehicle corresponds with one of the compass directions on the grid. In the third test, participants are presented with 12 road situations and are required to select the correct road sign for each road situation.

1.3 Procedure

Participants were initially asked about their driving experience and their previous experience of driving simulators. They were then seated in front of the VR driving simulator and the controls were explained to them. Each participant performed the practise route until they felt confident in their ability to steer the car, stop at the traffic lights, look both ways, and avoid other traffic. They then performed the actual driving assessment test. If they accidentally deviated from the route, the experimenter pressed a button which restored the car to the road again. On completion of the driving assessment test, each participant performed the four sub-tests of the SDSA. They were then thanked for their participation and the purpose of the study was explained to them.

3. RESULTS

No participants reported any adverse effects from using the driving simulator. All were able to complete the assessment and most reported that they enjoyed the experience. (Interestingly, some participants had reported simulator sickness effects in pilot studies when the driving simulator had been used with a very large screen.)

A measure of how well participants conformed to the speed limits of 30, 40 and 50 miles per hour imposed in the VR assessment was calculated by subtracting the relevant speed limit from participants' actual speed and converting all the resulting scores to positive. Sixteen of the 60 measures of stopping time were deleted because they were over ten seconds indicating that participants had, either slowed down without actually stopping, manoeuvred the virtual car to avoid a collision, or crashed. The remaining measures were divided by the speed at which the vehicle was travelling at the time and collapsed into one measure to reduce the number of statistical comparisons. Two measures of the time taken to merge with other traffic were also collapsed into one measure. Table 1 below shows the means and standard deviations of drivers' and non-drivers' scores in the VR driving assessment and the SDSA.

Table 1. *Drivers' and non-drivers' scores in the VR driving assessment and the SDSA*

Assessment	Measure	Drivers		Non-drivers	
		Mean	SD	Mean	SD
Hit car	No. of collisions with other cars	0.95	1.19	1.40	1.60
Hit curb	No. of collisions with curb	7.95	8.50	67.40	114.01
Steering	Degrees steering wheel turned	3267	984	3540	1250
Lane positioning	Distance from centre of lane (cms)	61.83	9.95	86.27	32.56
Stopping time	Time (secs) divided by vehicle speed	0.10	0.03	0.11	0.05
Total time taken	Time (secs)	622.62	102.08	641.07	70.09
Merge into traffic	"	30.97	9.21	30.70	10.34
Speed 30 mph	Non-conformance to limit (mph)	4.01	2.56	2.88	1.21
Speed 40 mph	"	5.05	4.33	6.71	3.38
Speed 60 mph	"	20.48	5.36	23.06	4.02
SDSA Dots	Time to complete (secs)	308.15	92.68	317.29	51.64
SDSA Compass	Cars correctly positioned	24.05	10.22	28.41	5.70
SDSA Road signs	Signs correctly positioned	7.85	2.83	6.76	2.91
SDSA Total score	In accordance with SDSA calculations	3.17	2.70	3.48	2.14

It would appear from Table 1 that drivers performed better than non-drivers in the VR driving assessment by having less collisions, hitting the curb less often, over-steering less, holding a better lane position, stopping marginally quicker, and conforming better to the 40 and 60 mph speed limits, However, drivers were not better than non-drivers at merging with other traffic or conforming to the 30 mph speed limit. Conversely, in the SDSA, drivers appear to have taken longer to complete the Dots Test, positioned less cars correctly in the

Compass Test, and achieved a lower overall score than non-drivers. The only measure in which drivers predictably appear to have scored higher than non-drivers is in the Road Signs Test.

Statistical analyses were performed to investigate whether there were any significant differences between drivers and non-drivers in the above measures. In all the statistical analyses, the probability level was set at 0.05. The only significant differences revealed by independent t-tests were the numbers of times the curb was hit [$t(38) = 2.33$, $p = 0.025$] and lane positioning [$t(38) = 3.21$, $p = 0.003$], with drivers performing better than non-drivers in both measures. None of the other comparisons reached statistical significance. The drivers' SDSA total scores significantly correlated in the predicted direction with the number of times that they hit the curb [Pearsons Correlation = -0.645, $p = 0.002$] and their over-steering [Pearsons Correlation = -0.480, $p = 0.046$] in the VR driving assessment. Other correlations between the SDSA total scores and the VR driving assessment measures did not reach significance.

4. DISCUSSION

Although drivers and non-drivers differed significantly on only two of the measures from the VR driving assessment, hitting the curb and lane positioning, the trend was that their performance was better than non-drivers in all but two of the measures. In contrast, the counter-intuitive trend in the SDSA total score was that non-drivers' performance was better than drivers'. Nevertheless, drivers' SDSA total scores significantly correlated with two measures from the VR driving assessment, hitting the curb and over-steering. It would appear that the cognitive tests in the SDSA are tapping into the same navigational skills which are used in steering the virtual car.

The two measures in the VR driving assessment where drivers' performance was not better than non-drivers' were "conforming to the 30 mph speed limit" and "merging with other traffic". The reason why drivers did not perform well in "conforming to the 30 mph speed limit" was that they were inclined to exceed the limit whereas non-drivers were not [drivers' mean = 30.89, non-drivers' mean = 29.95]. This tendency probably reflects their driving speeds in the real world. Unfortunately, the "merging with other traffic" measure did not work properly as it did not take account of participants simply turning left or right without paying any attention to oncoming traffic. Even if they crashed, they might still achieve a fast score on the "merging with other traffic" measure. This measure therefore has to be revised to take account of this possibility in the next version. Another measure which should be incorporated into a revised version of the VR driving assessment is a measure of reaction time which is independent of the speed of the car and avoidance actions by the driver, e.g. a notice appearing on the screen saying "Hit the brakes now". The current method of measuring stopping time in response to a hazardous road situation is too dependent on how the participant responds, which may be to steer around the hazard without stopping.

The finding that the VR driving assessment discriminated between drivers and non-drivers but the SDSA did not, indicates that only the VR driving assessment is measuring some of the driving skills acquired through experience. Research has shown that inexperienced drivers have problems with psychomotor (Mayhew & Simpson, 1996), perceptual (Mourant & Rockwell, 1972), and cognitive skills (Groeger & Brown, 1989). For example, inexperienced drivers are more prone to steering errors and inappropriate speed choices (Mayhew & Simpson, 1996). They also lack good hazard detection skills and take longer to respond to hazards (Groeger & Brown, 1989). The role of experience in driving cannot, therefore, be overestimated and any assessment that does not take driving experience into account when considering a person's ability to drive, is not providing a realistic measure of driving ability.

5. CONCLUSIONS

To many people, the ability to drive is an important aspect of their life and a symbol of their independence. However, after brain injury, or approaching old age, although people often feel that they are able to resume or continue driving, they are not always safe to do so. We have devised a VR-based driving assessment which will allow people to test their driving ability off-road and thereby help them to come to a more informed decision as to whether they should resume driving or continue to drive.

In the first stage of an evaluation of the VR driving assessment, drivers' and non-drivers' performance was compared on the driving simulator and the SDSA. The VR driving assessment discriminated between drivers and non-drivers on two measures but the SDSA did not, indicating that only the VR assessment is measuring actual driving experience. Since the role of experience is an important element in determining an individual's driving ability, a driving assessment test should take previous driving experience into account.

The results of the evaluation indicated that some measures from the VR assessment needed to be improved, particularly "merging with other traffic", and an independent reaction time measure should be incorporated into the next version. It will be worthwhile re-examining all the measures which did not produce significant differences between drivers and non-drivers to see if they can be improved.

After these revisions, and subject to two further stages in the evaluation, testing the VR driving assessment with people who have suffered a stroke, and comparing older people's performance on the assessment with their real-life driving ability, we are confident that this VR-based driving assessment will prove to be an invaluable aid to older people and people who have suffered physical or mental trauma in allowing them to decide themselves whether or not they should continue or resume driving. At the present time, the best off-road method of testing driving ability would be to use the VR driving assessment in conjunction with the SDSA.

6. REFERENCES

M P Barnes and E A Hoyle (1995), Driving assessment - a case of need, *Clinical Rehabilitation*, **9**, pp. 115-120.

British Psychological Society (1999), *Fitness to drive and cognition: a document of the Multi-Disciplinary Working Party on Acquired Neuropsychological Deficits and Fitness to Drive*, BPS, Leicester, UK

J A Groeger and I D Brown (1989), Assessing one's own and others driving ability: Influences of sex, age, and experience. *Accident Analysis and Prevention*, 21, pp. 155-168.

D R Mayhew and H.M. Simpson (1996). *The role of driving experience: Implications for the training and licensing of new drivers*. Insurance Bureau of Canada.

J M Mazaux, F Masson, H S Levin, P Alaoui, P Maurette and M Barat (1997), Long term neuropsychological outcome and loss of social autonomy after traumatic brain injury, *Archives of Physical Medicine and Rehabilitation*, **78**, pp. 1316-1320.

J Mollenkopf, F Marcellini, I Ruoppila, P Flaschentreger, C Gagliardi and L Spazzafumo (1997). Outdoor mobility and social relationships of elderly people, *Archives of Gerontology and Geriatrics*, **24**, pp. 295-310.

R R Mourant and T H Rockwell (1972), Strategies of visual search by novice and experienced driver, *Human Factors*, **14**, pp. 325-335.

F M Nouri and N B Lincoln (1993), Predicting driving performance after stroke, *British Medical Journal*, **307**, pp. 482-483.

F M Nouri and N B Lincoln (1994), *The Stroke Drivers Screening Assessment*, Nottingham, UK.

P Rabbit, A Carmichael, S Jones and C Holland (1996). *When and why older drivers give up driving*. Summary report. AA Foundation for Road Safety Research.

Design and evaluation of a flexible travel training environment for use in a supported employment setting

N Shopland[1], J Lewis[1], D J Brown[1] and K Dattani-Pitt[2]

[1]School of Computing & Technology, The Nottingham Trent University,
Burton Street, Nottingham. NG1 4BU, UK

[2]Learning Disability Services, London Borough of Sutton,
Mint House, 2nd Floor, 6 Stanley Park Road, Wallington. SM6 0EH, UK

nicholas.shopland@ntu.ac.uk, james.lewis@ntu.ac.uk, david.brown@ntu.ac.uk, kiran.pitt@sutton.gov.uk

[1]www.ntu.ac.uk, [1]www.isrg.org.uk, [2]www.sutton.gov.uk

ABSTRACT

This article describes the user centred design and development of a virtual environment (VE) to support the training of people with learning disabilities to travel independently. Three separate implementations were built on top of an initial design. Two of these environments implemented intelligent agents to scaffold learners using virtual environments; the third took stakeholder experiences to redesign the initial environment in an attempt to improve its utility.

1. INTRODUCTION

Many of the issues around travel training for people with learning disabilities (LaGrow 1990, Bourland 1996) and virtual reality (Brown 1997, Steed & Frecon 1999, Standen *et al* 2001) have been previously identified. Brown *et al* (2002[2]) report on a number of VEs developed to support people with learning disabilities in extending independent travel skills. Building on this experience, we describe the design and implementation of a three dimensional virtual environment for improving travel skills of people with learning disabilities, and on two approaches to incorporating virtual tutoring agents into such an environment. The travel training VE formed part of a larger "Invest to Save" project (LBS 2004) established by the London Borough of Sutton. Its aims are "to help adults with learning disabilities move into employment by training them in relevant subjects, offering support and helping them to find work".

Stakeholder meetings were held to set the initial design requirements for the VE, and from these meetings, storyboards were created and agreed upon. An initial environment was then built, implementing a limited sub-set of the stakeholder identified features in order to obtain feedback on a working design as soon as possible. This VEs development was then forked, allowing work to proceed in two areas:

1. An investigation of the application of virtual tutors as a means of scaffolding a trainee's experience within the 3D environment.
2. The development of a customisable environment enabling trainees to practice navigating a specific route and the hazards encountered along it.

Alongside the VE implementation, the Traverse project sought to incorporate handheld computers in support of its travel training programme. This was to be achieved by using *TouchSpeak*, the commercial result of a research project from the ACE Centre (2003), who developed a Portable Communication Assistant for People with Acquired Dysphasia (PCAD), a "hand-held augmentative & alternative communication device with a dynamic colour display, a built-in microphone and digitised speech output. It is designed specifically for non-speaking dysphasic and brain injured users".

2. PACKAGE DEVELOPMENT - PHASE 1

2.1 *Specification of the Package*

Before any design began, informal context meetings were held with trainers and service users to identify the elements important to the two principal stakeholder groups. These meetings sought to address three aspects of the design:

1. How the travel training software was likely to be used

2. What usability issues were likely to be encountered
3. Define an initial set of interactions

These issues were discussed separately with the trainers and also with trainers and learners in facilitated sessions. These sessions defined the initial content and design outline.

2.1.1 Intended Use of the Travel Training Software. The trainers identified their primary requirement for the software as to provide an environment where the users could practice real-world skills in a naturalistic manner. The trainer would then intervene where the learner exhibited unsafe behaviour or had difficulty in remembering the procedure for a particular task. This would allow the trainers to continue the principal method of their real world (RW) travel training procedures while working with the simulated environment.

2.1.2 Usability Issues. A technical limitation was in place at project inception; that implementation was to be using standard PC hardware. Given this constraint, the principal usability concern was the method of avatar control in the virtual environment. A variety of avatar viewpoints were planned and a range of control devices supported (keyboard input, mouse control and joystick control). On screen representation of control state was also considered as potentially useful. This variety of approaches was decided upon in order to test which were most usable and should be employed in subsequent development.

2.1.3 Initial Set of Interactions. The following elements were identified as an appropriate set of tasks that the environment should support.

Road Crossing	Public Transport
Normal road* Normal road with island Zebra crossing* Zebra crossing with island Pelican crossing* Pelican crossing with island One-way roads	Bus journeys, principally, catching the bus. Recognise a bus stop. Wait for a bus – look the correct way. Allow trainer to configure bus number (with destination?) and bus colour. Add possible decoy buses.

Additional aspects: More than one lane, Parked cars*, Straight & curved sections*

Only a sub-set of these elements (marked*) were selected for phase 1 implementation, namely, a selection of the road crossing interactions.

2.1.4 The Role of Touchspeak. Touchspeak was hoped to support the learner's real world communication in a similar manner to its use with dysphasic adults. There were also hopes that it could support learners in remembering the tasks and duties required of them by their employer.

2.2 Package Development

The phase 1 environment was given an urban setting, and included a user controlled avatar and traffic (parked and moving). A simple figure '8' road layout was developed, as it provided a simple environment in which a variety of road crossing elements could be implemented, and in which traffic could naturally flow. A park was included as a safe practice area to develop avatar control skills. Traffic density was variable and three crossing types were created at various locations around the scene. A number of predetermined tasks were incorporated into this environment, by defining start and end points for the avatar. A variety of control methods and viewpoints were included.

The three different control systems were:

- on screen buttons, allowing mouse or touch screen control
- keyboard control (using the arrow keys), also allowing for use of a simplified "concept" keyboard
- joystick control, also allowing the use of other analogue controllers such as a steering wheel

The three different viewpoints were:

- 1st person (avatar not visible) - Default
- 3rd person-fixed (avatar permanently in centre of view)
- 3rd person interpolating ("follow" camera)

Additionally, the cursor was necessary for a specific interaction (pelican button press), and so was changed to a large hand with two states, open hand (normal) & press (indicating an active area).

All configuration options were presented on the initial screen, and this screen was returned to whenever the simulation was exited. This screen offered 7 pre-set routes, plus an 8th manual option, with which start and end points could be set by the trainer. A cylindrical "hot area" identified the completion point of a task, this could be made visible (a yellow cylinder) or not in the virtual environment. Activation of the end point triggered a random congratulatory audio file. Settings for (parked and driving) traffic density were provided and the initial camera viewpoint.

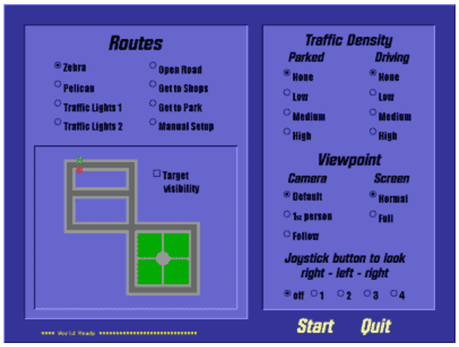


Figure1. *Phase 1 - the initial screen.*

Figure2. *Phase 1 in use, follow camera.*

2.3 Iterative Evaluation and Refinement

Once a working environment had been developed, user testing was conducted with a number of the client group trainees. This diagnostic evaluation identified a several of usability problems and potential areas for improvement.

Control of the avatar was found to be the main usability issue. Most users found that mouse control (by clicking on the directional buttons) was difficult because of the need for fine control and mouse clicking. Most users had much more success using the standard keyboard or a joystick. Turning was noted to be a problem with users frequently over steering. A suggestion from the trainers was to modify the method of avatar control by implementing "stepped" turning; this was implemented as a switchable option and improved the capacity of many of the trainees to control the avatar.

Looking to look up and down the street for oncoming traffic was found to be awkward since the avatar had to be turned to the left and right to bring the street into view. A camera action to look right, left, right was developed and bound to the space bar. This was also made available to be configured onto any of the first four joystick buttons.

A previous VR travel project (Lewis *et al* 2000) suggested that a first person perspective was an acceptable viewpoint for a similar group of users, so this was made the initial default. It very rapidly became apparent that the fixed 3rd person perspective was the preferred viewpoint, so the system default was changed to this.

An additional traffic density of zero was added, to allow for users to practice avatar control in the environment without any traffic. (The park was intended as a safe area to practice, but the trainers found that avatar control practiced in the park did not transfer too well to the street.)

Other improvements arose which were noted for consideration during the second development phase. The trainers were also interviewed to discover their experiences using the environment as part of their travel training programme and to determine those elements that they felt were important for inclusion in the second development phase.

3. INTELLIGENT AGENTS - PHASE 1

On completion of the first phase of development, two implementations of intelligent agents were developed for the environment (Chuang 2003, Penter 2003). These intelligent agents seek to scaffold the trainee's introduction to and use of the 3D environment. A baseline analysis of students and tutors (Chuang 2003) using the phase 1 environment without an agent suggested the principal activities of the tutor in supporting a user of the travel training software (and so those required of an agent) were to:

- Introduce the control devices
- Introduce the tasks
- Support route finding
- Provide problem solving whilst performing the tasks
- Monitor performance

These principal activities comfortably agree with the activities identified by the trainers at the package specification stage (2.1). This analysis of trainees using the environment helped to define the characteristics an Intelligent Agent requires; four characteristics were identified as being necessary (Chuang 2003):

- **Deductive Character –** The agent should introduce goals and learning objectives, and assist the trainee in mastering the control system.
- **Monitoring Character –** The agent should observe the student's progress, giving additional instruction, support and demonstration where the student is having difficulty or gets into a hazardous situation.
- **Motivational Character –** Successful achievement of goals should be rewarded with positive feedback, however, since some safety issues may have arisen from the student's behaviour despite a successful result, the agent may need to direct the student to repeat a task until no safety issues arise.
- **Solution-address Character –** The agent should provide incidental advice and guidance on request or in response to the breaking of a rule.

An interesting observation was that combination of two (successfully completed) simple tasks into a more complex task resulted in double the number of tutor interventions. Learners rapidly forgot the instruction previously given, and had to be prompted by the tutor; "[r]eminding users to do important tasks regarding safety issues should be included in IAs".

The Intelligent Agents were then implemented, incorporating these four characteristics, with two differing approaches, one using an active, animated 3D figure, and the other a static 2D representation of a tutor. Evaluation was conducted by observation, recording the duration and frequency of trainer interventions and user errors with and without a virtual tutor. Trainer assistance, task completion and user satisfaction were all noted to be enhanced in the environments augmented by an intelligent agent (Chuang 2003, Penter 2003). A set of design guidelines for the implementation of intelligent agents for people with learning disabilities were developed from these studies (Chuang 2003):

- Use simple and short language and avoid long sentences
- Consistent layout or controls for the Agent
- Use plain background for the speech balloon
- Use step by step instructions to guide users
- Provide positive feedback
- The speech of the text should be recorded by native speakers, or use text-to-speech techniques
- Provide dismissible functions while they are not essential
- Limit length of the animation
- Agent behaviour cannot be too complex
- Avoid use of voice input
- Minimize distraction from learning, for example give reminders without requiring learner's full attention

These guidelines build upon previous work by Brown *et al* (2001, 2002[1]) and Standen *et al* (2002).

4. FLEXIBLE TRAVEL TRAINING - PHASE 2

4.1 Refinement of Phase 1 Design

The design methodology planned for the project had always intended to undertake a radical overhaul of the prototype after the completion and evaluation of the work undertaken during phase 1. The rationale for this decision was that both the user group and the trainers, the stakeholders informing the design process, would be in a much better position to comment and suggest ideas after the experience of the phase 1 design cycle.

4.1.1 Stakeholder Input. The learners were very encouraging about the way in which they could practice tasks over and over again, building up confidence in their own abilities. There was a discernable view however that the lack of variety in the available tasks was making repeating the exercises a little monotonous. Users sometimes wanted to visit other parts of the environment, however because the crossing type they were learning to use was not located there they had to continue working in a specific area of the world. The trainers suggested that this could begin to effect the motivation of the users.

There was also a view from the trainers that the transfer of skills from the simulation to the real world could be made more reliable if the users were enabled to experience the same task in different contexts. It was suggested that by only exposing a user to a crossing type in one specific place, that its value as a generic skill may not be as effective as it could be. The trainers had fears that the skills required to understand how to use a crossing may be context specific e.g. *"this is what I do at the crossing outside McDonalds"* rather than the generic *"this is what I do at a pelican crossing"*. It was felt that if a particular task could be placed in different locations within the environment that the adoption of generic skills could more easily be achieved. The trainers suggested that the most important aspect to consider for this was the route following aspects of the users' experience. This was closely followed by the need for the users to associate parts of the journey with landmarks along the route.

A proposal was therefore made to overcome the limitations of the fixed VR world by developing a world builder with which the trainers could develop customised environments. This approach would allow the VE to reflect the sequence of actions and situations that a learner encounters along each of the routes that they are being trained on (as well as allowing the exploration of other, unfamiliar routes). The phase 2 design was therefore based on creating a set of 3D assets which could be combined into a world that went some way towards simulating the real world environment that the user would have to navigate their way around.

4.2 Implementation

The implementation of this idea was a tile based VE built from a variety of road tiles incorporating bends, crossroads, T-junctions, cul-de-sacs and straights, utilising both single and dual carriageways. The tiles were developed as a kit of interchangeable pieces which could be arranged to create a flexible customised environment that could be assembled to much more closely resemble the real world environment which the user would have to navigate their way through. Following the assembly of the road layout the trainer can populate the environment from a set of road safety interactions; zebra, pelican and puffin crossings can be dragged onto any road tile as can pieces of street furniture etc. This interface allowed any number of road and crossing configurations to be built by the trainer, enabling a varied presentation of various crossing types to be experienced by the user and also allowing the development of routes specific to the trainee's anticipated journey(s).

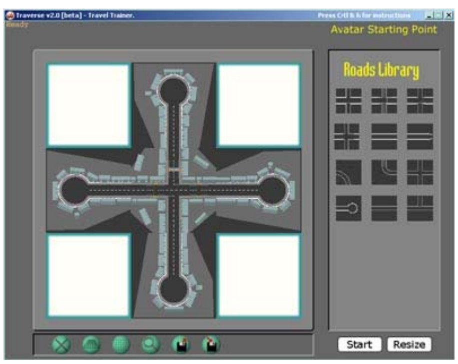


Figure 3. *The World Builder.*

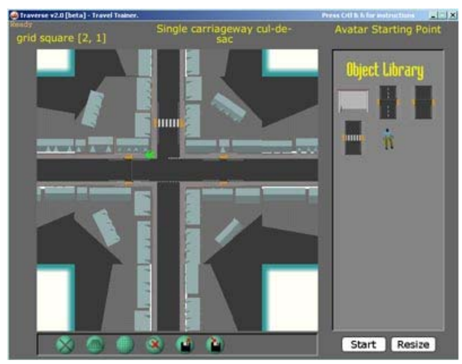


Figure 4. *Adding Interactions to the World*

4.2.1 Navigation Issues – Bridging the VE/RW Gap. The configurable environment seemed to enable generic skill acquisition and make a major contribution towards context specific route learning. However it was clear that the VE alone could not produce the infinite variety of real world situations, in which landmarks are features utilised by the users in navigating their way around. To tackle these issues the *Touchspeak* package was enlisted as a portable memory aid. The trainers together with the users surveyed the route noting the landmarks and decision points especially where they seemed to hold particular significance for the user. This survey was used to construct a hierarchy of photographic images and vocal and text prompts within *Touchspeak* that could be used by users throughout their journey, both in the simulation and in the real world.

At first the results of this were encouraging, however the trainers found the hierarchical nature of *Touchspeak* awkward to program and use, and the extremely flexible decision tree arrangement of the software was unnecessary for the linear progression of a journey. The approach was refined to use Microsoft PowerPoint to hold a slideshow of the same types of media, this time arranged in a linear fashion. The trainers found this approach much easier to develop for the users, and there was a suggestion that the users found this easier to use.

Whilst use of handheld computers seemed to offer much promise, several significant drawbacks were identified. There was some concern that the need to ensure that the units were kept charged could impact upon their usefulness - particularly where people with a learning disability took the units home. There were also concerns that somebody with a learning disability travelling alone with a handheld computer on prominent display could place themselves in a position where they were a target for theft. Indeed, the prominent display of valuable items such as handheld computers is contrary to police advice. Finally the user group expressed a view that in outdoor situations the display was very difficult to see - the problems of lack

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

of brightness and reflections from the LCD screen were exacerbated by the relatively small size of the display itself.

Following these trials the value of a portable memory aid was confirmed, the problem was perceived to be with the technology. After discussion with the users it was decided that a laminated paper based set of pictures and instructions was the most appropriate way of tackling the need for users to have reminders of decisions and actions whilst in the field.

4.2.2 Technical Refinement of the VE. Phase 2 offered the opportunity to solve a number of technical issues identified during the testing of the first phase. One of the most significant of these involved the complete reworking of the control system of the environment. Collision detection in the first environment was implemented through a system of ray casting. When this approach was first tested on the avatar it appeared to be a satisfactory solution, however once cars were introduced into the environment significant drawbacks emerged. The most critical instance of this problem was where a collision between avatar and car could go undetected. This was because ray casting was a frame-based test that became increasingly unreliable at the slower frame rates experienced on less powerful computers. This was exacerbated by the fact that the ray casting method could not detect partial collisions between geometry untested by the ray.

The solution to this was to implement the VE with a physics simulation. This simulation was used to monitor events within the environment and provide the control systems for both the avatar and the cars. Cars were given route finding and decision making routines, allowing them a degree of freedom to decide where in the world they would go, within the constraints of their rule set. To solve the collision issues experienced in phase 1, the physics simulation conducts collision tests 4 times between each frame draw, using low polygon versions of the VE geometry.

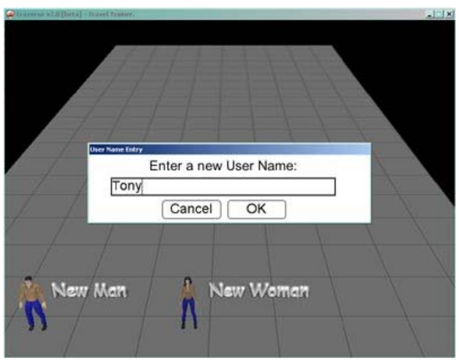


Figure 5. *The Lobby*

Figure 6. *Customising an Avatar*

The generic avatar used in phase 1 was always intended to be replaced, and discussion with the trainers and project management established that they required avatars to represent both sexes and a variety of races and ages. This was achieved by developing a male and a female avatar with customisable clothing, skin and hair colour. Implementing customisable avatars enabled the implementation of a simple log-in system. A 3D lobby was built, populated by the default male and female avatars (for new users) and any customised avatar used on the system. It was hoped to link this system up to user profiles containing personalised rotes and performance data. Observation of users demonstrated that they enjoyed being able to customise their avatar, a feature seeking to enhance identification of the trainee with the action taking place within the virtual environment.

A further refinement suggested by the trainers was to trigger puffin and pelican switches by walking the avatar into the control box. Implementing and testing this indicated that the former approach of using a hand cursor to represent pressing the crossing button was more useful, being more comprehensible to the learners and a more positive representation of the real world action. An additional usability issue was requested, to indicate when the avatar was correctly positioned at a zebra crossing - a visual cue was developed, however, we have reservations about using an on-screen indicator to give feedback to the user in this way, as such cues do not exist in the real world.

Figure 7. *The phase 2 VE*

5. CONCLUSIONS

The software was only one element within a general programme of life, employment and travel training skills. Because of this it is difficult to detach the contribution towards travel training made by the software, from the impact made by the trainers in the real world. Anecdotal evidence from the trainers suggests that the incorporation of V.R. training into a programme of travel training does have beneficial results. The skills learned in the V.R. environment appeared to transfer to the real world. There were indications that this skill transfer reduced the overall time taken to teach a user all of the skills required to enable them to travel independently.

Contrary to our initial expectations and implementation, the learners all preferred a 3rd person perspective for avatar control. Many users also found unmediated turning caused control problems; stepped turning appears to be a satisfactory method of addressing this.

Incidental outcomes identified by the trainers included a significant increase in the users confidence in using computer systems. There were also suggestions that the motor skills and co-ordination of users in operating mouse and keyboard were also improved.

Virtual tutors have been shown to improve the time taken to achieve a task in both of the tutor studies. However, it has not been demonstrated that such an improvement in one performance parameter associated with computer use leads to an equally effective transfer of skills from the VE training environment to the real world.

6. FUTURE WORK & LONGITUDINAL EVALUATION

The focus of this project was on the development of the travel training software following a user centred design methodology. Due to this, the focus of work was on the evaluation of the design from a usability and utility viewpoint. It is clear that there is a need for an extensive study to substantiate the anecdotal evidence of skill transfer to the real world and reduced overall training time.

There is also a particular need to look at the effects the software has on the users perceptions of hazards in the real world. This particularly applies if the software is to be used without of an observing trainer. A wider range of potential traffic hazards is planned for the simulation including emergency vehicles, reversing vehicles and parked cars moving off. A tendency to experiment with hazardous behaviour in the simulation has been noted, particularly where the user perceives this behaviour to have been "rewarded" because it generated a system response. There is a very real danger that inappropriate responses from the environment or an agent could desensitise users to real world hazards and it is essential to thoroughly evaluate this risk.

Work is ongoing to extend the range of tasks covered by the software, notably to incorporate the skills required to enable travel using buses and trains. There is also a desire to increase the set of tiles available to extend the range of environments and road shapes. This could enable urban, suburban and rural environments to be assembled to more closely match the needs of users. Extension of the user profiles to allow personalised routes and the collection of performance data is envisaged.

An additional area of interest raised by this project is the potential for exploring contemporary technology, notably 3G phones and GPS, as a means of augmenting travel training programmes. This is a wide open field that could use ideas such as a portable and usable mobile phone based slide-show tool to aid navigation or a bus timetabling system such as Nottingham's NextBUS (Bargiela 2004).

Funding for these activities is currently being sought. In the meantime a user survey questionnaire is being developed to give the trainers a subjective method of recording their opinions on how useful the travel training environment has been to their students. To widen the range of this feedback the software has been made freely available. It is available for download at http://traverse.isrg.org.uk. The user survey questionnaire will be circulated to users who have downloaded and registered their use of the software.

7. REFERENCES

ACE Centre (2003). PCAD TouchSpeak (Portable Communication Aid for Dysphasics) [online]. Oxford: ACE (Aiding Communication in Education) Centre Advisory Trust. Available at: http://www.ace-centre.org.uk/html/research/pcad/pcadproj.html [Accessed 14 July 2004].

Bourland, E. (1996). *Travel training for youth with disabilities.* Washington, DC: Academy for Educational Development, Inc., and National Information Center for Children and Youth with Disabilities.

Bargiela, A. (2004). ATTAIN - Advanced Traffic and Travel Information system [online]. Nottingham: The Nottingham Trent University. Available at: http://www.doc.ntu.ac.uk/RTTS/Projects/grr32468/attain.html [Accessed 15 July 2004].

Brown, D.J. (1997). The Virtual Reality Travel Training Package. *Internal Report, VIRART,* University of Nottingham.

Brown, D.J., Standen, P.J., Proctor, T. and Sterland, D. (2001). Advanced Design Methodologies for the Production of Virtual Learning Environments for Use by People with Learning Disabilities. *Presence: Teleoperators & Virtual Environments.* MIT Press. **10**(4): 401-415.

[1]Brown, D.J., Powell, H.M., Battersby, S., Lewis, J., Shopland, N. and Yazdanparast, M. (2002). Design Guidelines for interactive multimedia learning environments to promote social inclusion. In *Disability and Rehabilitation* (Brown and Rose Eds), **24**(11-12): 587-599.

[2]Brown, D.J., Shopland, N. and Lewis, J., (2002). Flexible and virtual travel training environments. In, *Proc. 4th Intl Conf. on Disability, Virtual Reality and Assoc. Technologies* (Sharkey, Sik Lányi & Standen, Eds), Veszprém, Hungary, 18-20 Sept. 2002, 181-188.

Chuang, W., (2003). Online Virtual Training Environments with Intelligent Agents to Promote Social Inclusion. M.Phil. thesis, Nottingham Trent University.

LaGrow, S., Wiener, W., & LaDuke, R. (1990). Independent travel for developmentally disabled persons: A comprehensive model of instruction. *Research in Developmental Disabilities*, **11**, 289-301.

LBS (2004). Traverse Project. [online] London Borough of Sutton. Available at: http://www.sutton.gov.uk/Sutton/Where+we+live/Transport/Traverse+Project.htm [Accessed 3 March 2004].

Lewis J., Brown, D.J. and Powell, H.M., (2000). Development of A Virtual Environment to Teach Independent Travel Skills to People with a Learning Disability. In *Proc. 12th European Simulation Symposium* (Möller, Ed), Hamburg, Germany, 28-30 Sept. 2000, 385-389.

Penter, L., (2003). *A Virtual Tutor for People with Disabilities*. B.Sc. dissertation, Nottingham Trent University.

Shopland, N., Lewis, J., Brown, D.J. and Powell, H.M., (2002). Virtual Travel Training for People with Learning Disabilities Accessing Employment Including the Introduction to the Special Thematic Session "Virtual Reality". In *Proc. 8th Intl. Conf. on Computers Helping People with Special Needs* (Miesenberger, Klaus & Zagler, Eds), Linz, Austria, 15-20 July 2002, 140-142.

Standen, P.J., Brown, D.J. and Cromby, J.J. (2001). The effective employment of virtual environments in the education and rehabilitation of students with intellectual disabilities. *The British Journal of Educational Technology.* **32**(3): 289-299.

Standen, P.J., Brown, D.J., Horan, M. and Proctor, T. (2002). How tutors assist adults with learning disabilities to use virtual environments. In *Disability and Rehabilitation* (Brown and Rose, Eds), **24**(11-12): 570-577.

Steed, A. and Frecon, E., (1999). Building and Supporting a Large-Scale Collaborative Virtual Environment, In *Proceedings of 6th UKVRSIG, University of Salford*, 59-69, 13th - 15th September 1999.

ICDVRAT 2004

Session III. Virtual Environments and Games for Special Needs

Chair: William Harwin

Computer games for children with visual impairments

Y Eriksson[1] and D Gärdenfors[2]

[1]Department of Art History and Visual Studies, Göteborg University,
Box 200, 405 30 Göteborg, SWEDEN

[2]Stockholm International Toy Research Centre, KTH,
Brinellvägen 85,100 44 Stockholm, SWEDEN

yvonne.eriksson@arthist.gu.se, dan@sitrec.kth.se

[1]www.hum.gu.se, [2]www.sitrec.kth.se

ABSTRACT

The Swedish Library of Talking Books and Braille (TPB) has published web-based computer games for children with different kinds of visual impairments. As the target groups have very different needs when it comes to the use of graphics and sound, TPB have developed two kinds of games. Image-based games aim to encourage children with partial sight to practise recognising visual objects, while sound-based games also intend to be accessible without relying on vision. Based on the results of two pilot studies, this paper discusses central design issues of the graphical and sound-based interfaces for this type of applications.

1. INTRODUCTION

Today, computer games are integrated parts of children's play activities as well as school education. Contemporary computer games usually feature sophisticated 3D-animated graphics, accompanied by sound effects to enhance the play experience. However, mainstream games are often inaccessible to children with visual impairments.

Children with partial sight frequently have trouble perceiving objects that constantly move around on the computer screen, and objects might be hard to distinguish because a lack of contrast in colours. Furthermore, the continuous noise of a soundtrack can be frustrating when trying to figure out the interface and the elements on the screen. For a blind person, it is practically impossible to play ordinary computer games, as they are based on visual output. To create accessible games for blind children, the games must be communicated by sound only.

For these reasons, TPB, the Swedish Library of Talking Books and Braille has published 13 computer games on their web site, targeted at children with different visual impairments (TPB 2003). The games are small Macromedia Flash™ applications, intended to serve as an introduction to computer games to people who previously have little experience of this type of entertainment. The games are published on the web to ensure that they can be accessed on various computer platforms, given that there is an Internet connection.

As children with partial sight and blind children have very different needs when it comes to the use of graphics and sound, the main aims have been (1) to design computer games that encourage children with partial sight to practise recognising visual objects, and (2) to create games that can be played without any graphics. Therefore, there are both picture-based and sound-based games, where the later category only requires aural attention to be played. The games with pictures are controlled with the mouse. The sound-based games are played with the computer keyboard, since many blind computer users avoid the mouse. The games have intentionally been designed for home computer equipment, so one only needs a standard PC or Mac with a pair of headphones or loudspeakers to play the games.

2. ACCESS TO IMAGES

Many children with partial sight are not very interested in using their sight because of the big efforts and the small rewards involved. Since they easily can misinterpret visual representations, they generally expect little useful information from pictures. The graphics of traditional computer games usually contain a lot of details to add realism or to make the games more attractive, such as 3D effects and colour nuances. This can distract children with partial sight, since they very easily lose their focus and have difficulties finding the main characters in the games. Therefore, a computer game for people with partial sight has to be designed with clear and simple illustrations, where excessive details are avoided. The TPB games also include slow movements to stimulate the players to look at the graphics.

For a person with limited sight, it can be difficult to get an overview of a computer game situation. With software such as LunarPlus (Dolphin Oceanic 2004) or ZoomText (Ai Squared 2004), computer users can enlarge what is displayed on the screen and read text at a size that they find comfortable. While these applications are convenient for reading, they are not as suitable for watching pictures and especially not animated graphics, since only a small part of the screen can be seen at a time. The TPB games contain illustrations that are created according to the target-groups special needs for high contrast and clear figures. Therefore, the games use vector-based graphics, which also means that the player image can modify the image sizes. The illustrations only show the distinguished features of the objects and characters, so they do not necessarily look realistic in the same way as photographs do.

2.1 Shapes

When we see an object and identify it as, for example, a vase, a shoe or a dog, it is often because we can recognise its shape. The shape of an object is found in its outline. Outlines do not exist in reality although our brains are constituted so we automatically see contour lines that make it possible to organise our visual perception. This is one of the reasons why we can distinguish objects in the foreground from the background, both in the environment and in visual representations. According to Arnheim (1966), perception of shapes is the understanding of structural features found in, or imposed upon, the stimulus material. When we depict an object, we draw the outlines of it and if the representation is successful it can be identified by others. Even very simplified drawings of an object or a person are easy to interpret if the artist has caught its significant shapes, which is something we can experience in caricatures as well as silhouette cut outs.

The TPB games for children with partial sight feature high-contrast pictures with designs inspired by Gestalt psychology theories. Influenced by Koffka (1934), Eriksson (1995, 1997, 1998, 1999) stresses the importance of outline shape. The shape of an object is dependent of the angle from which it is observed, which is especially noticeable in representations. As shown by Solso (1994), it is much harder to identify a represented object if it is depicted from an unusual angle. For an observer to recognise an object from its outline shape, a representation has to be made from an angle that makes the actual object identifiable. For people with visual impairments it is often hard to interpret tactile as well as visual representations of objects that diverge from the physical form, since the contour of the object will appear very differently from the real object. Therefore, the graphics of the TPB computer games feature objects and characters that are depicted either in profile, *en face* or from above, to facilitate interpretation.

2.2 Colours

Colour plays a considerable role in contemporary visual media. However, when interpreting visual representations, colours are subordinated shape. We identify a picture of a dog because we can recognise it from the significant shapes of its different features, its nose, its body and its legs. It does not matter if the dog is red, green or brown. Furthermore, if we see a figure in the same colour we normally associate with a specific dog, we will not interpret the representation as a dog if its shape is similar to a cat.

Together with light and shadows, colours are often used to create plasticity in paintings or drawings. By manipulating different nuances of a colour, it is possible to add volume to an object on a two-dimensional surface. This has been used by numerous artists and is often applied in mainstream computer games. Eriksson (1999) emphasises that for a person with visual impairment, the colour modulations can cause problems, since the parts with other nuances easily can be interpreted as separate objects. The same issue arises when using shadows and light in images. Fluidised colours can also be problematic, since their luminary quality can obscure the contours of objects. The quality of the contour is crucial for recognising the outline of a shape. The graphics of the TPB games use neither colour nuances nor shadows. The pictures are designed to look flat, so the contour of every object is distinct from the background. In most games, the background is dark, while the objects have light colours. This colour scheme seems to be the most efficient, since it prevents different picture elements from appearing as silhouettes. While a dark silhouette clearly indicates the outline shape, it can prevent a person with limited sight to perceive details within the figure. A dark background makes the player perceive that the backdrop settles while the objects move forward. This is ideal for a person with visual impairment, especially for someone who lacks colour vision.

The notion of contrast between colours usually refers to complementary colours, meaning those found on the opposite sides of the colour circle, such as orange and Klein blue. Contrasts that depend on complementary colours have the same degree of saturation and therefore they can be difficult to perceive for people with colour blindness. For a person who perceives the environment in a grey scale, there is no contrast between colours with the same degree of saturation. As colour blindness is a common effect of visual impairment, the TPB games are made in colour contrast that is based on different levels of saturation.

2.3 Image-based games

The TPB games intended for children with partial sight include jigsaw puzzle games, memory games, action games and play environments where the main objective is exploring.

There are three jigsaw puzzles with different themes. One puzzle starts by briefly showing the player the final picture of a horse before the player can start moving the pieces. The puzzles featuring a snake and a castle do not start by showing the motifs, so the player has to discover the themes by putting the parts together step by step. The puzzle with a castle shows a building with two towers (see Figure 1). While towers are significant features that one generally associates with castles, the motif of the jigsaw puzzle is not understandable for children who lack the knowledge of how to interpret these symbols. Essentially, it is necessary to have knowledge of a specific symbol to be able to interpret its meaning. A clear layout of a picture does not necessary implement an easy interpretation. As discussed by Eriksson and Göthlund (2004) among others, even a picture that appears as very simple is often quite arbitrary.

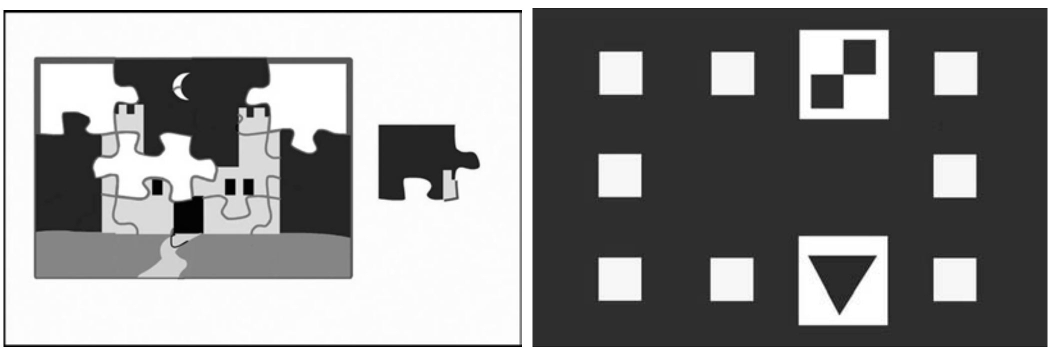


Figure 1. *The castle jigsaw puzzle and a memory game. Graphics by Maria Beskow, courtesy of TPB [TPB 2003].*

The four memory games have different themes, such as farm animals or abstract moving objects. Two of the memory games feature cards with different symbols, such as semi-circles, circles, triangles and squares (see Figure 1). One game has objects with light colours on a dark background, another has dark figures on a light background. The variations in designs are chosen to offer different challenges to the player.

The *Beetle game* requires fast input from the player who tries to stop beetles from eating a cake. The main challenge is to synchronise the movement of the mouse with the movement of the beetle across the screen, as it requires simultaneous awareness of two moving objects.

3. SOUND-BASED GAMES

Most mainstream computer games feature high quality sound effects and soundtracks that enhance the visual animations. However, as sounds mainly are added as embellishments, they do not convey enough information for blind players to be able to understand what happens in the game. Since the sounds cannot be clearly connected to a specific position on the screen, visual information is required to understand where they come from. Furthermore, there often are many characters, objects or events that are not associated with any sound at all.

To make computer games accessible to blind people, one cannot simply add more sounds to existing computer games. Friberg and Gärdenfors (2004) stress the importance of planning the entire game concept around an auditory experience in order to develop a comprehensive sound-based game. Since sounds are very different to graphics with regard to space and time, sound-based game development requires a very different design approach to that of graphical games. In the TPB sound-based games, all the game interfaces, the menus as well as the gameplay, are designed from an entirely auditory perspective.

The sound-based games belong to two categories, the puzzle games *Towers of Hanoi* and *Memory* and the action games *Tag* and *Skybells*. In addition to auditory interfaces, these games feature the same type of high contrast graphics as is used in the other TPB games (see Figure 2). By adding graphics to complement the audio, which is the opposite procedure compared to mainstream game development, the games ultimately feature two complete interfaces. Thus, blind children can play together with friends with partial or full sight, who can choose to use the graphics if they find the sound-based gameplay too difficult.

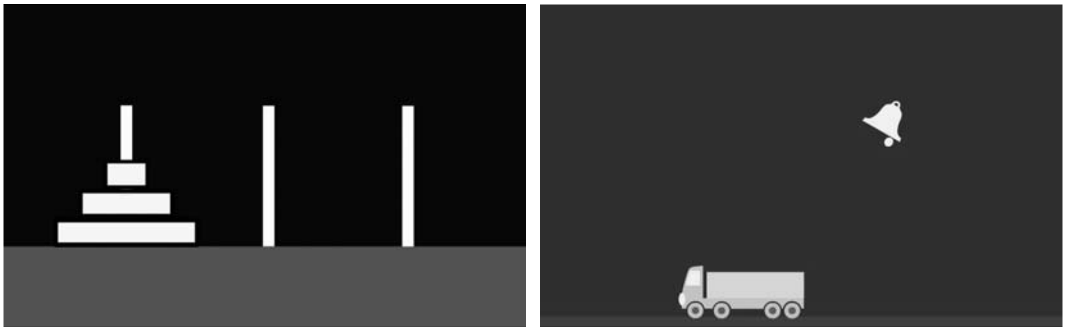


Figure 2. *Graphics from Towers of Hanoi and Skybells. Graphics by Maria Beskow and Annica Norberg, courtesy of TPB [TPB 2003].*

3.1 Positioning sounds

In sound-based games, the player obtains a mental image of all present objects and characters by listening to the sounds that they are associated with. To separate different sound objects spatially, the TPB games rely on stereo positioning, which enables sounds to be spread from left to right. This information is essential for the players' understanding of the game setting. However, as stereo only represents one dimension, it offers a limited space compared to the two dimensions of the computer screen. Just like one visually can simulate depth by graphical trickery, it is possible to imitate more spatial dimensions sonically by head related transfer function (HRTF) techniques. As this requires very advanced sound engines, it is currently not an option when developing web-based games. According to Menshikov (2003), HRTF needs headphones to work well. While headphones can be convenient when playing the TPB games, they are not made obligatory since people with visual impairments might not want to be entirely shut off from the outside world.

The TPB games rely on less realistic ways to convey depth and height. In the Tag game, an illusion of depth is created by alterations in sound pressure. Height is indicated using two different strategies. In *Memory* and *Skybells*, the objects that are higher up are represented by sounds with higher pitches than those of objects below. The *Towers of Hanoi* game relies on a different convention. When checking a horizontal position in this game, the present objects are presented from the bottom and up, so only one object is heard at a time. Obviously, these types of conventions take some time for the players to learn. As it is harder to convincingly communicate spatial relationships with sounds than with images, sound-based games can benefit from being more stylised, with their own, non-realistic game logic.

3.2 Auditory overview

To avoid excessive auditory information, all the TPB sound-based games are designed to stage very few objects, with simple spatial relationships. In the fast-paced games *Tag* and *Skybells*, the game objects are in constant motion, generating continuous sounds. However, in the puzzle games, *Memory* and *Towers of Hanoi*, the game objects are still unless the player activates them, which causes a more static gameplay. In these games, it would be very difficult to distinguish the different objects if they were to emit continuous sounds.

Instead of continuous display, the TPB puzzle games require that the player checks the different positions in the game field to hear the object sounds. Therefore, the sounds always appear in sequence, which relieves the player of having to take in too much simultaneous information. This means that the player must memorise the brick positions when playing the game, or frequently check the different positions. Winberg and Hellström (2000) have created an auditory Towers of Hanoi application that features continuous sounds. While this approach provides a constant overview to the player, the uninterrupted sounds blend into a complex code, which is something the TPB game aims to avoid.

3.3 The sonic palette

Relying on sounds to communicate all aspects of a game, from the menus and instructions to the actual game objects, many types of information need to be presented by different auditory interfaces. There are several kinds of sounds that can be used to convey different types of messages, such as speech, music or "sound effects." The TPB sound-based games use all these kinds of sounds to indicate objects, events and continuous processes within the game environments.

Recorded or synthesised speech is useful when communicating very precise information, such as game instructions and menus. However, inside the actual games, the use of speech is limited, since it generally is too slow to communicate events that occur in a high tempo. Furthermore, spoken messages tend to grow tedious if repeated too often. Neither is speech suitable if more than one message is conveyed simultaneously.

The auditory interfaces of the TPB games are mainly made up from "sound effects," which here refers to sounds that are intended to correspond to the graphics of visual games. These illustrative sounds can be designed in several ways. When possible, it can be useful to represent objects by authentic sound recordings, as they are often easy to recognise. However, since many objects and events do not generate sounds in the real world, several sounds have to be designed from other sources. As it is convenient to design interfaces that build on users' previous experiences, Gaver (1986) argues that "iconised" sounds can be useful in auditory interfaces. "Auditory icons" are sounds that somehow relate to the objects that they represented in the interface. This design approach is used in the TPB *Towers of Hanoi* game, where stone discs make "clinking" sounds, and the wooden poles sound "woody." Like graphics, sounds can be stylised to different degrees. When moving a disc sideways, one hears somewhat cartoonised "swish" sound, which is intended to give the impression of an object being moved, even though stone discs rarely make such sounds.

Another approach to auditory interface design (Brewster et al, 1992) is to use abstract musical sounds to create auditory messages. This type of auditory symbols, also called "earcons," show little or no resemblance to what they represent. This means that earcons can be used to indicate any phenomenon, generating pleasant, musical interfaces. The main drawback of earcons is that they can take a long time to learn, as they are more arbitrarily linked to the real world than "auditory icons." In the *Memory* game, a type of "earcon" created by pitched percussion instruments is used to represent different positions in a grid of cards. These sounds were chosen because the action of checking positions is not clearly associated with a sound in the real world. Other examples of events that are musically illustrated in *Memory* are the rewarding sounds played when finding a pair of cards or winning the game, and the error alert sound heard when attempting an impossible move

There are endless possibilities to combine the auditory interface design methods above. One reason for this is that it is possible to listen to a sound in many ways. Influenced by Schaeffer (1966), Chion (1994) divides human listening into three modes: casual listening, semantic listening and reduced listening. Casual listening applies when listening to the source of a sound, attempting to understand what caused it. Semantic listening is used when understanding auditory codes such as speech or Morse code. Reduced listening is used when listening to qualities of a sound without considering its source, such as when appreciating music by listening to its pitches, harmonies and rhythms.

As Chion (1994) claims that it is possible to apply more than one listening mode simultaneously, it is possible to design sounds that are both iconic and musical the same time. Friberg and Gärdenfors (2004) suggest a system that illustrates the various auditory interface design approaches according to Chion's three listening modes. In analogy to a system used to describe visual cartoons by McCloud (1993), a triangular model can be used to indicate the different ways in which sounds can be designed (see Figure 3).

This triangular sonic palette demonstrates the range of different sound design strategies that can be used when developing auditory interfaces. The TPB games are not limited to speech, the "auditory icon" or "earcon" design approaches. Instead, they use a sonic palette that spans between voice, authentic sounds and musical sounds.

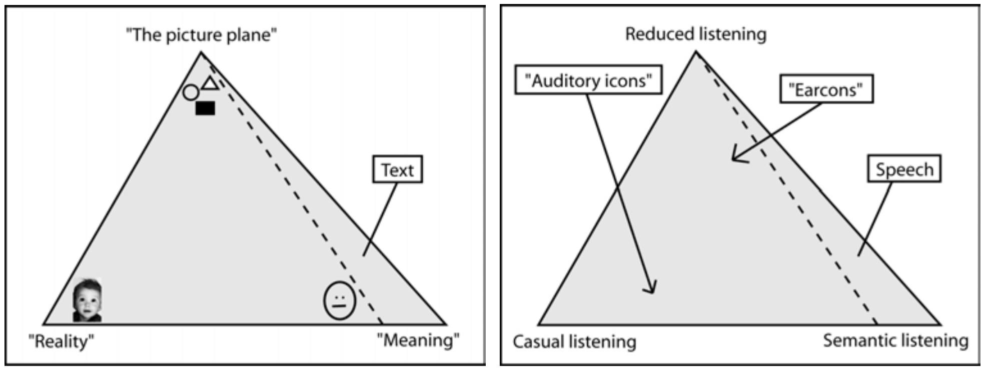


Figure 3. *The caption McCloud's triangle of visual representation and a corresponding triangle for sonic representation. Graphics courtesy of Friberg and Gärdenfors [2004], after McCloud [1993].*

4. TEST RESULTS

The TPB game project has involved two pilot studies: with five children, between 6 and 12 years old in 2001, and four children between 9 and 12 years old in 2003. The first test focused on the image-based games. To be able to analyse how and if these games work for a child with limited sight, we made an eye tracking study. All the children in the group have heavy nystagmus, a constant, uncontrolled eye movement that is very common among people with partial sight. It has been discussed if nystagmus leads to limitations in the ability to interpret visual information, especially when it is represented by images. However, we found that the children had no problem playing the computer games and that they could coordinate the movement of the mouse with what they saw on the screen. This was clear in the *Beetle Game*, in which it is important to react quickly with the mouse to smash a beetle before it takes a piece from the cake.

Surprisingly, the children found the horse jigsaw puzzle harder to play than the *Beetle Game*. They said that the time that the motif of the puzzle was displayed before the game starts was too short. To solve this task methodically, and not by chance, one has to remember the outline of the horse, and one must be able to associate the different details on the pieces with a specific part of the total image. This can be difficult for children with visual impairments, since they often only can observe a limited part of the screen at a time.

The second user study showed that blind children found the sound-based *Memory* difficult to play. Three out of four children could not manage to figure out the organisation of the cards on the screen. The one who succeeded had played the game earlier together with her father and, with his help, created a mental map of the surface.

Skybells, like *Memory*, is space related in the sense that the player has to understand that the bells and the stones fall in four different horizontal positions. Some children had difficulties both in perceiving the stereo positioning and in understanding how to avoid the stones and to catch the bells. The children found the game much easer to play when using headphones instead of loudspeakers, which indicates that headphones are more reliable when conveying differences in stereo positioning.

One child with limited sight, who after a careful explanation could play Skybells quite well, found it much more difficult when he was asked to use his sight. As his sight is very limited, and since he first learned the game from its sounds, he got confused when exposed to the visual and auditory information simultaneously. In this specific case the pictures and the sounds did not complement one another for the child.

Only one child in the second pilot study was totally blind. So, as a second trial we asked the children to play the games using their sight. One of the girls, 12 years old and classified as blind with an almost non-existing sight, made a better result with the sound-based *Memory* when using the graphics.

The *Towers of Hanoi* game proved to be a great challenge for the children, especially those who do not have access to the visual representation of the objects and have to rely on sound only. To be able to solve the problem, the player has to understand the spatial relationship between the objects before it is possible to start moving the discs. No child under 9 years has been able to complete the *Towers of Hanoi*. As the game rules are quite sophisticated, all children need careful instructions before being able to solve the game.

Tag is an action game that requires skill when controlling the game with the keyboard. One of the children in the second pilot study had a motor problem that lead to difficulties in managing the keys. Still, he found the game very stimulating and it will work as a platform for practising and challenging the limitations of his motor skills.

5. CONCLUSIONS

The game project by TPB has shown that it is possible to create fun computer games for children with partial sight or blindness. In general, the response to the TPB games has been very positive, which indicates that several of the design hypotheses may be adequate. The fact that most children with partial sight, nystagmus and even one child classified as blind could play the image-based games signifies that the visual designs of the TPB games are successful.

Still, several issues remain unsolved or little explored, especially regarding the sound-based games. It is difficult to communicate static overviews of several objects with sounds. Sound, just like text, is limited when it comes to conveying unchanging environments. Visual game interfaces offer the player a continuous overview that makes it possible to investigate environments at random. Pictures are also exceptional when it comes to conveying spatial relationships between different objects.

The pilot study indicates that, when it comes to sound-based games, it is easier to play action games such as *Tag* than a static puzzle, such as *Memory*. This is probably due to the lack of overview offered in the sound-based puzzle games. Therefore, it is very important to provide clear instructions for games that require a spatial understanding to be solved. For younger children, instructions have to be introduced by a person. It is not enough with a pre-recorded verbal instruction that accompanies the game.

The game elements developed in the TPB project can be used to develop larger and more advanced computer games for people with visual impairments. Several ideas have already been used successfully in other game projects, such as TiM (Friberg and Gärdenfors 2004). Still, the main challenges when creating sound-based games remain, mainly regarding how to design interfaces that efficiently utilise the spatial and temporal characteristics of sound. Headphones seem to be advantageous when understanding the positioning of sounds, however they cannot be the only output option since they shut off the player from the outside world.

It seems that even children with very limited eyesight intuitively find it easier to interpret visual images than auditory representations. This might be because they, like most people, more often are exposed to conventional visual symbols than to non-speech auditory interfaces. When designing computer games that rely on non-speech sound, there is a lack of conventions on which the auditory interfaces can be based. While Western culture has a rich tradition of visual iconography, there is no well-established auditory counterpart. Without an established canon of auditory icons, players of audio games today need elaborate instructions before they can start playing.

Acknowledgements: The authors would like to thank TPB for funding the game development. Maria Beskow initiated the game project together with Yvonne Eriksson. SITREC, KTH and the Department of Art History and Visual Studies, Göteborg University, supported the research.

6. REFERENCES

Ai Squared (2004), http://www.aisquared.com/

R Arnheim (1966), *Visual Thinking*, California University Press, Berkeley.

S A Brewster, P C Wright and A D N Edwards (1992), A detailed investigation into the effectiveness of earcons, In G. Kramer (Ed.), *Auditory display, sonification, audification and auditory interfaces. The Proceedings of the First Intl. Conf. on Auditory Display,* Addison-Wesley, Santa Fe, pp. 471-498.

M Chion (1994), *Audio-Vision: Sound on Screen*, C Gorbman Ed., Columbia University Press, New York.

Dolphin Oceanic Ltd (2004), http://www.dolphinuk.co.uk/

Y Eriksson and E Tebelius-Murén (1995), On making tactile: an exhibition project in Nationalmuseum, *Art Bulletin of Nationalmuseum Stockholm*, Vol 1-2.

Y Eriksson (1997), *Att känna bilder*, SIH Läromedel, Solna.

Y Eriksson (1998), What some people see is what others feel, *Art Bulletin of Nationalmuseum Stockholm*, Vol 5.

Y Eriksson (1999), "Jag vill ha en hund". Multimedia för barn med synskada, *Bilder ocn internet - texter kring konstruktion och tolkning av digitala bilder, visuellt*, Konst- och bildvetenskapliga institutionens skriftserie nr 6.

Y Eriksson and A Göthlund (2004), *Möte med bilder: Att tolka visuella uttryck*, Studentlitteratur, Lund.

J Friberg and D Gärdenfors (2004), Audio games: New perspectives on game audio, *Proc. ACM SIGCHI Intl. Conf. Advancements in Computer Entertainment Tech.,* Singapore, pp. 148-154.

W W Gaver (1986), Auditory icons: Using sound in computer interfaces, *Human-Computer Interaction*, **2**, *2*, pp. 167-177.

K Koffka (1935), *Principles of Gestalt Psychology*, Kegan Paul, Trench Trubner & Co. Ltd., London.

S McCloud (1993), *Understanding Comics: The Invisible Art*, HarperCollins, New York.

A Menshikov (2003), Modern audio technologies in games, *Game Developers Conference*, Moscow, http://www.digit-life.com/articles2/sound-technology/index.html

P Schaeffer (1966), *Traité des Objects Musicaux*, Seuil, Paris.

R L Solso (1994), *Cognition and the Visual Art*, MIT Press, Cambridge Mass.

TPB (2004), Web-based computer games and written report, http://www.tpb.se/english/games

F Winberg and S-O Hellstöm (2000), The quest for auditory direct manipulation: the sonified Towers of Hanoi, *Proc. 3rd Intl Conf. Disability, Virtual Reality & Assoc. Tech.,* pp 75-81.

Virtual reality rehabilitation for all: Vivid GX versus Sony PlayStation II EyeToy

D Rand[1,2], R Kizony[1] and P L Weiss[1]

[1]Department of Occupational Therapy,
University of Haifa, ISRAEL

[2]Beit-Rivka Geriatric Medical Center,
Petach-Tikva, ISRAEL

[1,2]*drand@univ.haifa.act.il*, [1]*rkizony@univ.haifa.ac.il* [1]*tamar@research.haifa.ac.il*

hw.haifa.ac.il/occupa/LIRT

ABSTRACT

The main objective of this paper was to investigate the potential of the Sony PlayStation II EyeToy (www.EyeToy.com) for use in during the rehabilitation of elderly people with disabilities. This system is a projected, video-capture system which was developed as a gaming environment for children. As compared to other virtual reality systems such as VividGroup's Gesture Xtreme (GX) VR (www.vividgroup.com), the EyeToy is sold commercially at a relatively low cost. This paper presents three pilot studies which were carried out in order to provide essential information of the EyeToy's potential for use in rehabilitation. The first study included the testing of healthy, young adult participants (N=18) and compared their experiences using the EyeToy system to the GX system in terms of sense of presence, sense of enjoyment, control, success and perceived exertion. The second study assessed the usability of the EyeToy with healthy elderly subjects (N=10) and the third study assessed the use of the EyeToy with stroke patients (N=8). The implications of these three studies are discussed.

1. INTRODUCTION

Clinicians who work in rehabilitation aim to enhance clients' functional ability as well as their ability to participate in community life. These goals are achieved by intensive intervention aimed at improving motor, cognitive and meta-cognitive abilities. For many injuries and disabilities the rehabilitation process is very long and arduous, and clinicians face the challenge of finding appealing and motivating intervention tools that will facilitate this process. Virtual Reality-based therapy appears to provide an answer to this challenge due to its well-known assets including the opportunity for experiential, active learning, the ability to objectively measure behaviour in challenging but safe and ecologically-valid environments, while maintaining strict experimental control over stimulus delivery and measurement, and the capacity to individualize treatment needs, while gradually increasing the complexity of tasks and decreasing the support provided by the clinician (Schultheis et al., 2001; Rizzo, 2003).

Although the advantages of VR are becoming widely recognized within the clinical community, the rehabilitation team faces a daunting challenge – to find an off-the-shelf VR system that enables achievement of the goals stated above, yet affordable by the typical clinical facility. A still greater challenge is to find motivating intervention tools that a client could afford to acquire for home-based therapy. This last point is particularly important since, in recent years, it has been demonstrated that only intensive repetition of exercise leads to significant improvement in functional ability (Liepert et al., 2000) yet, due to the high cost of in-patient hospitalization, rehabilitation centres are forced to reduce the time a patient remains in hospital. Thus, therapeutic exercise, initiated and monitored within the occupational therapy or physiotherapy departments, is insufficient on its own, and clients should be provided with opportunities to exercise at home.

For the past few years, a number of clinical research groups have explored the therapeutic potential of VividGroup's Gesture Xtreme (GX) VR (www.vividgroup.com). GX VR is a projected video-capture system; participants stand or sit in a demarcated area viewing a large monitor that displays a environment or functional tasks, such as touching virtual balls. A single camera, vision-based tracking system captures and

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

converts the user's movements for processing; the user's live, on-screen video image corresponds in real time to his movements, leading to engagement in the simulated task.

This system was originally developed for entertainment purposes, and has been adapted for use in rehabilitation (Kizony et al., 2003; IREX - www.irexonline.com). Different researchers have recently investigated this system (Kizony et al., 2002; Reid, 2002; Sveistrup et al., 2003) demonstrating its suitability for use during the rehabilitation of patients suffering from motor and/or cognitive deficits. Its advantages include the fact that patients see themselves rather than being represented as an avatar. They do not have to wear special apparatus such as an HMD which encourages the use of active movement and reduces their chances of experiencing side-effects. In addition, the therapist can intervene easily during the session in order to support and guide the patient's movements (Kizony et al., 2003).

Despite these many advantages, GX VR is still not widely used in rehabilitation facilities. One of the major reasons for this limited usage is the cost of the system that keeps it out of the price range of many clinical settings. The PlayStation II EyeToy (www.EyeToy.com), recently released by Sony, Inc, is an off-the-shelf, low-cost gaming application, which provides the opportunity to interact with virtual objects that can be displayed on a standard TV monitor. As with the Gesture Xtreme system, the EyeToy displays real-time images of the user. However, it does not require a chroma key blue/green backdrop behind the user nor bright ambient lighting. This makes for an easier setup of the system in any location but, on the other hand, it means that the user sees himself manipulating virtual objects within a video image of his own physical surrounding rather than within different virtual environments. The EyeToy application includes many motivating and competitive environments which could be played by one or more players (e.g., boxing, spinning plates) as well as different visual effects which encourage active movement without giving a score (e.g., painting a rainbow, mirror image distortions and popping bubbles).

The EyeToy's low cost and easy setup are advantages that have encouraged us to consider its use in rehabilitation, and to investigate its usability for wide segments of the population and not just for children who were the original market targeted by this application. It is also essential to verify that the levels of exertion evoked by this application are within the physiological capacity of patient populations, many of whom are elderly or weak.

2. OBJECTIVES

The main goal of this study was to assess the potential of the EyeToy for use in rehabilitation with people who are elderly and have disabilities, such as those who have had a stroke. This was achieved by performing three different pilot studies:

Study 1: To compare the GX and EyeToy applications using healthy young participants in terms of their effect on users' sense of presence, level of enjoyment, perceived exertion and side effects.

Study 2: To assess the usability of the EyeToy with an elderly healthy population in terms of sense of presence, level of enjoyment, perceived exertion and side effects and their ability to operate the system.

Study 3: To assess the use of the EyeToy with patients who have had a stroke in terms of their ability to cooperate, use different applications, and their level of enjoyment and perceived exertion.

3. STUDY 1

3.1 Participants

Eighteen healthy participants aged 21-37 (mean age 25.3 ± 4.0 years), mostly university students, volunteered to participate in the study. To date, 17 subjects were females and one was male. This ongoing study will eventually include more male participants.

3.2 Instruments

3.2.1 Virtual Reality Systems and Environments. VividGroup GX System was used with two virtual environments; Birds & Balls - touching balls that approach the user in a pastoral setting (Fig. 1a) and Soccer - preventing balls from entering the goal crease (Fig 1b). The Sony PlayStation EyeToy was used with two virtual environments too; Wishy-Washy - cleaning windows (Fig 1c) and Kung-Foo - fighting off other Kung-Foo fighters who are attacking (Fig 1d).

*3.2.2 Presence Questionnaire (PQ)***.** The PQ was translated from Witmer & Singer (1998) and used to assess presence. It is composed of 19 questions in which participants use a 7-point scale to rate various experiences within the VE; the maximum total score is 133 points indicating a high level of presence. The items assessed different aspects of presence: involvement/control, natural, interface quality and resolution. This was administered after each system.

Based on Slater's (2003) recent comments about the construct of presence, an additional question was asked: "During the time of the experience, which was strongest on the whole, your sense of being in the virtual environment, or of being in the real world of the laboratory?" This question was rated on a scale from 1 (being in the real world of the laboratory) to 7 (being in the virtual environment). This question was asked after participants experienced each system.

3.2.3 Scenario Feedback Questionnaire (SFQ). The SFQ is based, in part, on a translated version of Witmer and Singer's (1998) Presence Questionnaire) and was administered after each environment (virtual game). These six items assessed the participant's (1) feeling of enjoyment, (2) sense of being in the environment, (3) success, (4) control, (5) perception of the environment as being realistic and (6) whether the feedback from the computer was understandable. Responses to all questions were rated on a scale of 1-5, which were combined to give a global response to the experience for a maximum score of 30. An additional question was added to inquire if the participants felt any discomfort during the experience.

3.2.4 Borg's Scale of Perceived Exertion (Borg, 1990). This scale was used to assess how much physical effort the participants perceived that they expended during each VR experience. This is a 20-point scale that participants rated from 6 (no exertion at all) to 20 (maximal exertion).

3.2.5 Performance. Performance was monitored using the scores obtained in each environment (environment). For the Wishy-Washy environment the number of windows in addition to the total number of pints was recorded and for the Kung-Foo, the number of enemies and total number of points was documentation.

3.3 Procedure

Each participant experienced two virtual environments for three minutes on each of the VR systems: For the GX, Birds & Balls and Soccer and for the EyeToy, Wishy-Washy and Kung Foo were used. For all four environments the users were required to move their arms and bodies in order to interact. After experiencing each virtual environment, the participants were asked to complete the Scenario Feedback Questionnaire (SFQ). After experiencing the two environments using the first VR system, Witmer and Singer's10 Presence Questionnaire was completed. Participants then underwent the same procedure with the second VR system. The order of VR systems was balanced across participants. The entire experimental procedure took place during a single session lasting about 40 minutes. At the end of the procedure the participants were asked to rate the four environments in terms of enjoyment.

3.4 Data Analysis

Paired t-tests were used in order to determine differences between the sense of presence for each of the two systems (as assessed by the PQ and Slater's (2003) Presence question). Repeated measures were used to assess the differences between the four virtual environments (for the total SFQ, for the first question of the SFQ (sense of enjoyment) and for perceived exertion). This was followed by paired t-tests in order to identify the source of the significance.

3.5 Results

As a first step for each analysis we examined whether the order of experiencing the VR systems influenced the results. There were no significant differences due to the order in which the VR systems were experienced by participants for any of the outcome measures.

Figure 1. *Screen shots of the virtual environments. The two GX environments included (a) Birds & Balls and (b) Soccer. The three EyeToy environments included (c) Wishy-Washy, (d) Kung-Foo and (e) Keep-Ups.*

3.5.1 Differences between systems. The mean total PQ score for the GX environments was 94.1 ± 9.3 points (out of a maximum 133 points) and for the EyeToy environments was 93.7 ± 8.4. These mean scores were not found to be significantly different. The mean score for Slater's presence question for the GX system was 3.6 ± 1.2 points (maximum 7 points which indicates full presence in the virtual environment); this was not significantly different than the mean of 4.1 ± 1.6 points for the EyeToy system.

3.5.2 Differences between the four environments experienced by each participant. Significant differences between the environments were found for the total SFQ ($F(15)=5.970$, $p<.007$). The score for Kung-Foo was 25.3 ± 3.4 points which was significantly higher than the scores for Soccer (21 ± 3.6 points) ($t(17)=-3.88$, $p<.01$) and for Birds & Balls (21.8 ± 3.2 points) ($t(17)=-3.96$, $p<.001$) (see Table 1).

Significant differences were also found for the sense of enjoyment (the first question of the SFQ) ($F(15)= 12.06$, $p<.00$). When paired t-tests between the different environments were used, significant differences were found between Kung-Foo and Birds & Balls ($t(17)=-5.02$, $p<.000$), between Kung-Foo and Soccer ($t(17)= -4.507$, $p<.000$) and between Kung-Foo and Wishy-Washy ($t(17)=-3.716$, $p<.002$) all in favour of Kung-Foo.

The differences in perceived exertion between each of the four environments were found to be significant ($F(15)=12.068$, $p<.000$), where Birds & Balls was considered to be significantly easier than Soccer ($t(17)=-6.460$, $p<.000$), Wishy-Washy ($t(17)=-3.401$, $p<.003$) and Kung-Foo ($t(17)=-3.987$, $p<.001$) (see Table 1).

4. STUDY 2

4.1 Participants

Ten healthy elderly participants (six women and four men) aged 59-80 (mean age 70 ± 5.7) years were included.

4.2 Instruments

4.2.1 Virtual Reality Systems and Environments. Sony PlayStation EyeToy was used with three virtual environments; Wishy-Washy - cleaning windows, Keep-ups - bouncing a virtual soccer ball (See Fig. 1e) and Kung-Foo - fighting off other Kung-Foo fighters who are attacking

4.2.2 Scenario Feedback Questionnaire (SFQ). This scale ranges from 6-30 points (See above for further details)

4.2.3 Borg's Scale of Perceived Exertion [9] *(See above for further details)*

4.2.4 Usability Questionnaire. This questionnaire is composed of ten questions; five positive and five negative items regarding the use of the system. Each system was rated from 1 (disagree totally) to 5 (agree totally). The absolute sum of the 10 questions was calculated for a total score which ranged from 10 to 50 points.

Table 1: *Results from the total SFQ, sense of enjoyment and Borg scale for the four virtual games.*

	GX System		EyeToy system	
	Soccer	**Birds & Balls**	**Kung-Foo**	**Wishy-Washy**
Total SFQ	21 ± 3.6	21.8 ± 3.2	25.3 ± 3.4	23.5 ± 4.4
Sense of enjoyment	3.6 ± 0.92	3.6 ± 0.85	4.7 ± 0.43	4.2 ± 0.92
Borg scale (6-20)	13.6 ± 2.4	10.6 ± 1.8	13.4 ± 2.8	12.9 ± 2.6

4.3 Procedure

The participants experienced the EyeToy in their homes. Prior to the two minute experience with each environment, they were given a minute to practice. After each environment they were asked to fill in the SFQ and rate their perceived exertion. Their performance within each environment was recorded. Upon completion of the three environments, the participants were requested to independently exit the Kung-Foo environment and to start the Keep-Ups environment. The sequence of their actions was monitored. They were then asked to fill in a usability questionnaire regarding the use of the EyeToy.

4.4 Data Analysis

To assess differences between the three virtual environments (for the total SFQ, the first question of the SFQ – sense of enjoyment) and perceived exertion repeated measures were used, followed by paired t-tests. Independent t-tests were then used to compare the results of the ten elderly participants to the 18 younger participants in terms of their SFQ, perceived exertion and performance in two of the environments (Wishy-Washy and Kung-Foo).

4.5 Results

Significant differences were found between the environments for the participants' sense of being in the environment, of enjoyment, control and success as reflected from the total SFQ score ($F(8.0)=7.724$, $p<.014$). The SFQ score for the Wishy-Washy environment was the highest (25.3 ± 3) and it was found to be significantly higher ($t(9)=-3.245$, $p<.01$) than Keep-Ups (18 ± 5.7). The SFQ score for Kung-Foo (23.2 ± 4.5) was also significantly higher than Keep-Ups ($t(9)=4.134$, $p<.003$).

Regarding the sense of enjoyment (the first question of the SFQ), significant differences were found between the three virtual environments ($F(8)=5.024$, $p<.04$); paired t-test revealed that the significant differences were between Keep-Ups (3.5 ± 1.0 points) and Kung-Foo (4.5 ± 0.7 points) ($t(9)=-3.354$, $p<.008$); Wishy-Washy (4.3 ± 0.9 points) did not differ significantly from the other environments. The perceived exertion for all three environments was very similar, 11.0 ± 2.6 for Kung-Foo, 11.3 ± 1.4 for Wishy-Washy and 11.3 ± 1.4 for Keep Ups. No significant differences were found. A mean usability score of 43.3 ± 4.0 points (maximum possible score = 50) was found as a result of the participants operating the system on their own.

The overall experience using the EyeToy system by the younger (N=18) and older (N=10) participants was compared. No significant differences between the groups were found for either Wishy-Washy or Kung-Foo (see Tables 2 and 3 for further details). During the Wishy-Washy environment, participants from both groups succeeded in cleaning a mean of 5 windows during the 180 s environment. The younger group reported feeling more fatigue upon completion of the environment than did the older group but this difference was not significant.

During the Kung-Foo game, no significant differences for the total SFQ were found between the groups, however, significant differences were found for the Borg scale, and the number of enemies killed during the first game. The Kung-Foo game is not limited by time but you have two lives before "game over". In order to reach a full two minutes of engagement in the game, some participants had to continue playing into a second game. Most of the young participants managed to complete one full game, taking two minutes to do so. In contrast, the older participants needed 1.9 ± 1.1 games on average (range 1 to 4 games) to complete the two minute experience. This difference was significant ($t = -2.586$, $p<.03$).

Table 2: *Differences between age groups for the Wishy-Washy environment*

	Age group 59-80 N= 10	Age group 21-37 N=18	t	p
SFQ (6-30)	25.3 ± 3	23.5 ± 4.4		NS
Borg scale (6-20)	11.3 ± 1.4	13 ± 2.6		NS
# window's cleaned	5.5 ± 2.3	5.6 ± 2		NS
Total points	15,124 ± 7,523	15,888 ± 8,117		NS

Table 3: *Differences between age groups for the Kung-Foo environment*

	Age group 59-80 N= 10	Age group 21-37 N=18	t	p
SFQ (6-30)	23.2 ± 4.5	25.3 ± 3.4		NS
Borg scale (6-20)	11 ± 2.6	13.4 ± 2.8	2.147	.041
# of enemies killed	46 ± 40	86 ± 52	2.108	.045
Total points	13,780 ± 12,997	23,033 ± 16,137		NS

5. STUDY 3

Due to the positive response of the healthy elderly participants in Study 2, it was considered appropriate to test the EyeToy with participants with stroke.

5.1 Participants

Seven patients with stroke participated in this study. Two of the participants lived at home and had a mean post-stroke interval of 2.5 years. Five patients had a stroke within the past 2.5 months and were still in the process of receiving acute rehabilitation. All seven patients had sustained a right hemispheric stroke causing left hemiparesis. Table 4 presents additional data on each of the seven patients.

5.2 Instruments

The Sony PlayStation EyeToy was used with Wishy-Washy (cleaning windows) and Kung-Foo (fighting off other Kung-Foo fighters who are attacking). The Scenario Feedback Questionnaire (SFQ) rated from 6-30 points and Borg's Scale of Perceived Exertion were also used (see above for further details).

5.3 *Procedure*

The two chronic patients experienced the two virtual environments at their home while the five patients in acute rehabilitation experienced the environments at the hospital. They were given a practice minute prior each of the environments, followed by 2 minute experience using first Wishy-Washy and then Kung-Foo. After each environment they were asked to fill in the SFQ and rate their perceived exertion.

5.4 Data Analysis

Due to the small sample size of this initial study with stroke patients, descriptive statistics were used when available. In some cases, it was only possible to report on our observations of performance and reaction to the EyeToy.

5.5 Results

5.5.1 Chronic stroke patients. The two chronic stroke patients reported great enjoyment during their experiences with the EyeToy games. During the Wishy-Washy game they both requested to hold a towel in their hand, a prop that appeared to help them perform the task of cleaning windows. They all enjoyed the game very much (5/5). The total SFQ for both participants was 27/30 points. They rated exertion at 13-15 points and reported fatigue of their weak upper extremity. They were encouraged to use their weaker left arm as well as their strong arm and they did succeed in use it for cleaning the left side of the window. The number of windows cleaned was 1 or 2 with 1100-2988 points. This is much less than the scores of the healthy elderly although we cannot do statistics due to the small sample size.

Both participants enjoyed Kung-Foo very much (5/5 points) with a total SFQ of 27-28 points; they preferred using it more than Wishy-Washy. The level of exertion was rated in the same manner as for Wishy-Washy (13 to 15 points). Their Kung-Foo character was eliminated very quickly causing the game to end; this took 18 to 30 sec for the first game, 12 to 45 sec for the second game and 36 to 48 for the third game. These fast eliminations did not appear to frustrate them since they understood its therapeutic value (especially its value as good exercise for their weak upper extremity). Moreover, they were aware of and encouraged by the fact that their scores improved with each successive game; managed; they gained 400 to 2,100 points for the first game, 1,600 to 6,400 for the second and 3,400 to 6,300 for the third.

5.5.2 Acute stroke patients. All five patients enjoyed their experience very much and said they would happily repeat it. Several of the participants had difficulty using the SFQ to rate their answers. Participants 4 and 5 became frustrated during the games since they could not use their weak upper extremity to interact with the virtual objects. They understood the therapeutic value for their upper extremity and therefore were not satisfied when they accomplished the task using their right unaffected arm. We noted that all patients had difficulty in interacting in the space beyond their own bodies. Rather, they primarily moved their hands in front of their body. This was a limitation, especially for the Wishy-Washy game since the current window must be completely cleaned before the next one appeared. Many of the patients would have been stick cleaning the same window without help from the therapist. Participant 2 was relatively young and had considerable active movement in her left weak arm. Out of the five acute patients, she was the one who gained the most from the experience with the EyeToy.

Table 4: *Characteristics of the patients with stroke*

	Age	*Sex*	*Months post stroke*	*Functional status – independence*	*MMSE*	*Mobility*	*FMA- Left upper extremity*
1	*91*	*F*	*2*	*dependant*	*24/27*	*wheelchair bound*	*37/60*
2	*59*	*F*	*1*	*BADL - partial*	*20/28*	*wheelchair bound*	*45/60*
3	*72*	*F*	*2*	*dependant*	*NA- Aphasia*	*wheelchair bound*	*52/60*
4	*73*	*F*	*3*	*dependant*	*27/30*	*wheelchair bound*	*11/60*
5	*70*	*M*	*2*	*dependant*	*28/30*	*wheelchair bound*	*18/60*
6	*74*	*M*	*30*	*BADL & IADL*	*30/30*	*walk independently*	*48/60*
7	*69*	*M*	*36*	*BADL & IADL*	*30/30*	*walk independently*	*32/60*

F-female, M- male
FMA – The Fugl-Meyer Motor Assessment assesses the ability to perform different movements with the weak upper extremity. The scores ranges from 0 (no active movement) to 60 (full active movement).
BADL- Basic activity of daily living (such as eating, dressing, bathing).
IADL- Instrumental activities of daily living (such as cooking, shopping, house chores).
MMSE – Mini Mental Status Examination- cognitive screening test (maximum 30 points).
NA- Not applicable due to Aphasia (language disorder).

6. SUMMARY, CONCLUSIONS & RECOMMENDATIONS

The main goal of this study was to assess the potential of the EyeToy for use in rehabilitation with people who are elderly and have disabilities. Three pilot studies were carried out in order investigate the EyeToy's potential. The first comparison study demonstrated that the young participants sensed the same level of presence while experiencing the EyeToy as they had while experiencing the expensive GX system. Overall the participants enjoyed the four environments, but had a clear preference for Kung-Foo. The exertion levels were the highest for Soccer and the lowest for Birds and Balls. The importance of second pilot study was the fact that the elderly healthy participants enjoyed using the EyeToy, especially during the Wishy-Washy followed by Kung-Foo and then the Keep-Ups game. A high score for the usability questionnaire was obtained, indicating their satisfaction with the system and its ease of use. The EyeToy system, and especially the Kung-Foo, was found to be sensitive and differentiate between young and older participants and between stroke patients.

Assessing the use of the EyeToy with stroke patients emphasized the EyeToy's limitation, in particular the inability to grade the level of the environments. The EyeToy seemed to be less suitable for the acute stroke patients since they suffered from severe weakness of their left side of their body accompanied by sensory, cognitive and language deficits in some cases. Due to this fact, some of the patients expressed frustration especially when they could not manage to interact with the images with their weaker hand. Due to this limitation, the therapist is often required to help patients, for example, by bringing them closer to the camera (to make them appear larger) or by seating them on one side instead of the middle in order to facilitate performance. In some cases it is necessary to provide the patient with physical guidance or help. Despite this fact all of the acute stroke patients enjoyed their experience and expressed interest in repeating the session. However, the games appear to have the most potential for chronic patients or patients who sustained a mild stroke. Another limitation of the EyeToy is that the data recording is not sufficient .

The assets of the EyeToy, which were demonstrated by the results of all three studies, are that it is low cost, easy for users to operate, interesting, motivating and enjoyable. There is no doubt that these are all very valuable as an intervention tool during the rehabilitation of stroke patients and those with other neurological disorders. Motivated patients would be encouraged to practice movements in a repetitive manner using the EyeToy, thereby improving their condition. This is not easy to achieve via conventional therapy (Liepert et al., 2000). Moreover, the EyeToy appears to be ideal for use by the patient at home, where their healthy elderly caregiver would be able to operate the EyeToy for them. This is very important since the rehabilitation process after stroke is very long and, indeed, never ends. Patients living at home can benefit from using the EyeToy which requires the use of active movement of the whole body, attention, and rapid responses.

It is nevertheless unfortunate that despite the promise of using the EyeToy with a patient population, these environments cannot be graded to suit low functioning patients or in order to train specific therapeutic goals. During the past several years, other, similar products came on the market (e.g., by Intel and Reality Fusion) but were discontinued. Given the need for a low-cost VR video-capture tool that can be graded it is hoped that new products will soon become available.

7. REFERENCES

G Borg (1990), Psychophysical scaling with applications in physical work and the perception of exertion, *The Scand. J. Work Environ. Health, 16 Suppl 1*, pp. 55-58.

R Kizony, N Katz, H Weingarden and P L Weiss (2002), Immersion without encumbrance: adapting a virtual reality system for the rehabilitation of individuals with stroke and spinal cord injury, *Proc.4th Intl.l Con. Disabil. Virtual Reality Assoc. Technol.* University of Reading: Vresprem, Hungary, pp 55-61.

R Kizony, N Katz and P L Weiss (2003). Adapting an immersive virtual reality system for rehabilitation. *J.Visual. Comp. Anim. 14*, pp. 261-268.

J Liepert, H Baunder, H.R Wolfgang, W.H Miltner, E Taub, & C Weiller (2000), Treatment-induced cortical reorganization after stroke in humans. *Stroke, 31*, 1210-1216.

D.T. Reid (2002), Benefits of a virtual play rehabilitation environment for children with cerebral palsy on perceptions of self-efficacy: A pilot study. *Pediatric Rehabilitation*, 5, 141-148.

A.A. Rizzo (2003). A SWOT Analysis of the Field of Virtual Rehabilitation. Keynote address at the Second Intl Workshop on Virtual Rehabilitation. Piscataway, NJ, USA: www.caip.rutgers.edu/vrlab/iwvr/2003

M Slater (2003), A Note on Presence Terminology. *Presence- connect, 3, 3*

M.T. Schultheis, & A.A. Rizzo (2001). The application of virtual reality technology for rehabilitation, *Rehab Psychol*, 46, 296-311.

H Sveistrup, J McComas, M Thornton, S Marshall, H Finestone, A McCormick, K Babulic and A Mayhew (2003), Experimental Studies of Virtual Reality-Delivered Compared to Conventional Exercise Programs for Rehabilitation. *CyberPsychol. Behav., 6,* pp. 245 – 249.

B.G Witmer, & M.J Singer (1998), Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*, 225-240.

Game accessibility case study: Terraformers – a real-time 3D graphic game

T Westin

Department of Education, Stockholm University,
Utbildningsvägen 3, Tumba, Stockholm, SWEDEN

thomas.westin@mmedu.net

www.mmedu.net

ABSTRACT

Terraformers is the result of three years of practical research in developing a real-time 3D graphic game accessible for blind and low vision gamers as well as full sighted gamers. This presentation focus on the sound interface and how it relates to the 3D graphic world, and also include post mortem survey results from gamers and comments to those.

1. INTRODUCTION

Terraformers is the result of three years of practical research in developing a real-time 3D graphic game accessible for blind and low vision gamers as well as full sighted gamers. The idea when we started the project was to try bridge the gap in the game industry between the mass market 3D games for full sighted and the sound only games for blind and low vision gamers. The game was released on December 18, 2003 and can be found at www.terraformers.nu.

The game won the "Innovation in Audio Award" at the 2003 Independent Games Festival (www.igf.com). This in turn gave me the inspiration to start up the Game Accessibility Special Interest Group at IGDA (see Further work below).

2. ACCESSIBILITY FEATURES

2.1 Overview

Developing a 3D game demands in itself technical innovation in many areas like 3D graphics and other media optimization, game physics, game AI and so on. However, in this presentation I will focus on the sound interface and how it relates to the graphic interface, since this combination is the most original aspect of Terraformers.

2.2 Technical innovation and original, practical implementation

The game is a visual / audio hybrid playable by both sighted and sight disabled / blind. The following features describe how the 3D world is accessible to sight disabled. Some features allow direct access to the 3D world through use of 3D sound (i.e. spatial sound), other features allow access through voice feedback. Also, all game objects are accessible in a similar way (not only the 3D world space).

Powersuit: the powersuit contains all the necessary technology to protect the avatar from hostile attacks, In addition, it provides all of the accessibility features described below. A built-in voiced PDA is responsible for miscellaneous feedback from these features.

Sound Compass: a 3D sound represents north, and a rough 8 direction spoken feedback is available by pressing a key on the numerical keyboard (north, northwest...).

Sonar: a 3D sound gives the gamer a rough perception of the distance to objects in the direction the gamer is currently facing. By pressing a key the gamer can also check what type of object it is (door, wall, robot...). Enemies are automatically told by the PDA voice.

GPS: a global positioning system is used to get the exact positions of objects in an area as well as the position of the avatar. A voiced menu system provides an overview of nearby objects.

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

Direct orientation: using the numeric keyboard the gamer can orient the avatar directly in 8 directions (north, northeast...)

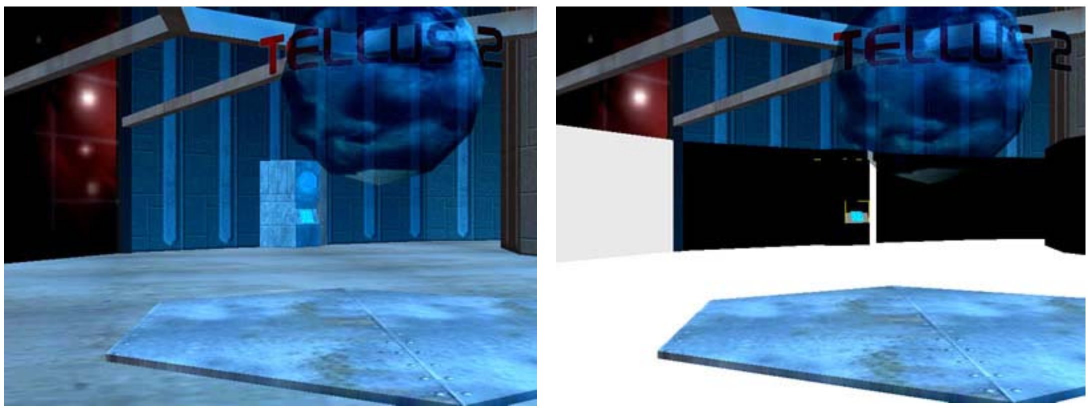
High Contrast Mode: With this feature turned on, you get parts of the 3D graphics rendered in black and white toon-style which enhance contrast on important objects for low vision gamers. See Figure 1.

No 3D Graphics Mode: The 3D graphic rendering can be completely turned off for blind gamers who don't have 3D graphic hardware installed. There is also a possibility to set the 3D engine to not use 3D graphic hardware, i.e. it the use the regular CPU only which is a way to workaround malfunctioning graphic drivers in Windows (which some blind gamers don't want to bother about).

Backpack: objects are accessed using a voiced menu system with hierarchies. Sound effects can also be attached to voices dynamically.

Game objects: all game objects have voiced feedback and 3D sound icons (and 3D graphics!)

Environments: 3D graphics, ambient sounds, footstep sounds on different ground materials and voiced descriptions of visual as well as other sensory input contribute to environmental feedback.



3D scene rendered in normal contrast for sighted gamers

3D scene with important parts rendered in high contrast, for low vision gamers

Figure 1. *Normal and High Contrast Real-time 3D Rendering.*

3. USER EVALUATION

3.1 User evaluation during development

This was performed ad hoc during the development process. We found it hard to do user evaluation since there was no other 3D graphic game that had a sound interface for blind when we started the project in November 2000. We had to first develop something, and then test it in an iterative process..

We did testing on several occasions with an adult blind programmer and kids in the age of about 13-16 years at a school for blind in Stockholm. We also got quite extensive feedback from a worldwide audience after we released the first official demos of the game, and we changed and added things that was possible to implement.

Since this has been a practical research with focus on development rather than documentation we don't have much of these evaluations in written form. Rather they are represented by the interface in the actual game, which is described in some detail above and also available for download at the game website (www.terraformers.nu).

3.2 Post mortem user evaluation results

The data was collected by submitting an e-mail survey to the first 35 customers, i.e. after the development of the game was finished (post mortem). The answers were open ended and allowed for respondents to write whatever they liked.

Table 1. *Survey results.*

	Blind, 49	***Blind, 11***	***Blind, 45***	***Blind, 19***
#1 Overall usability	*Very easy*	*Easy*	-	*I think that the game is well done and the usability is also good*
#2 Is the Sonar easy to use?	*Very usable*	*Easy to use*	*Yes*	*It's easy to use. I haven't had any problems.*
#3 Is the Compass easy to use?	*Very easy*	*Easy to use*	*Yes*	*It's OK*
#4 Is the GPS easy to use?	*Yes*	*Easy to use*	*Yes*	*This is also OK*
#5 Is the Backpack easy to use?	*Yes*	*Easy to use*	*Yes*	*This is OK too*
#6 Is the tutorial sufficient for learning the interface?	*Yes but I feel that it should cover all of the game* 2)	*Yes it is*	*Yes*	*It is sufficient*
#7 Is the game too easy / too hard?	*At the moment I can't answer that because I need to play the game properly* 3)	*The game is a little too easy* 4)	*No*	*It's not too easy and not too hard.*
#8 Are there interface solutions that you would like to see in a future release?	*I don't know at the moment* 5)	*No*	-	*Nothing at the moment*
#9 What do you think of the game as a whole? What is best, what is worst?	*I like it and when I have played it properly then I will be able to say more about it* 6)	*What I think is best is the ending of the game, and what I think is worst is the way the game responds and the character walks too slow* 7)	*Very good over all. Sometimes you run out of energy, I need more time before needing to recharge.*	*It's hard to answer to this question... I haven't found any bad things to say but the crashes that drives me crazy :) I think that there could be also some more levels.* ***8)***
#10 Are the enemies and quests too hard / too easy? If yes, why?	*again at the moment I can't say* 6)	1)	*No*	*They are very challenging*
#11 Have you completed the game? If yes, in how many hours approximately? If no, how many levels have you finished (approximately)?	*I am still on the first level* 6)	*Yes I have completed the game and it took me 1 hour and fifty minutes*	*Yes. Didn't count the hours. I completed the game within 3 weeks of buying it. When can I hope to get another game from you, this was fun.*	*I have completed the game approximately 2 months ago. It took about 8 hours to complete the game.*

We got 4 replies that was filled out according to the form. To these some minor edits have been applied (mainly spell correction) to make it more easy to read. A fifth reply was also received. It didn't follow the form so it was hard to compare it with the other replies. We will do a follow-up on that reply later.

As seen in the matrix in Table 1 there are numbers referring to comments and questions raised when interpreting the replies.

Comments on the matrix:

1) The reply for this was so long that I found it easier to read outside of the matrix:

   What I think about the enemies is that they are _way too easy, because all you really do with them is turn and face them and shoot them. I know they use electric weapons of their own, but I think their to slow and they can only reach you at a distance where when they aren't that close enough to use their weapon you can easily shoot them. and shouldn't they have something other than weapons too? For example maybe that life form at the end of the game it could move to the side when you first shoot it and come after you more quickly and maybe have a really long tail or something that would kill you if you got hit by it and something that would take a lot of health away from you if you did get hit by that particular body part. Could put a little more of robots and different kinds of enemies in the game? I think the quests are just right.

2) I'm not sure what the respondent refer to here. Virtually all in-game interfaces are covered by the tutorial, except for some special things but they are well documented. Of course the optimal thing would be to have everything covered in the tutorial but we have focused on what was found to be most important. I have not been able to follow up on this issue yet.

3) I have not been able to follow up on this issue yet.

4) It would be interesting to know more here but I have not been able to follow up on this issue yet.

5) I have not been able to follow up on this issue yet.

6) I have not been able to follow up on this issue yet.

7) It would be interesting to know exactly what the respondent refer to with "the way the game responds and the character walks too slow". This could be a technical issue with performance on his particular computer, or it could be a design issue. I have not been able to follow up on this issue yet. The game has been tested to run OK (although not optimal) on a Pentium II, 350 MHz with a GeForce II graphic card and SoundBlaster live soundcard.

8) It would be interesting to know more of the possible causes to crashes. Real-time 3D games require stable graphic and sound drivers (since the graphics and sounds are processed at runtime). This is a common problem for all 3D game developers. However, for Terraformers this poses a special problem: How do you motivate blind gamers to update their graphic drivers? On the other hand, the graphics are motivated for gamers with low vision (running in high-contrast mode) and of course for sighted friends and family. We have been working hard to provide information and help about how to fix issues like this. (More information can be found at the support section at www.terraformers.nu) I have not been able to follow up on the issue with this respondent.

4. CONCLUSIONS

4.1 Summary of the survey and discussion

Since the post mortem study is based on four replies only, we can't make any general conclusion about how the game interface works in it's final state based on the survey. On the other hand, during the three years of development we have had many more users trying out the game and giving feedback along the way, which is represented by the game itself. Hence I think you should read this report *and* play the game with sound only to get a better understanding of how the game works for sight disabled. I don't see the point in describing the game in more detail with text since there is a fully playable demo at www.terraformers.nu .

I'll just sum up what has been described in earlier sections regarding the game and the game interface:

- The game was considered good overall
- The game's overall usability was interface was easy to use (question #2-5). The tutorial was also considered good enough for learning the interface (question #6).
- All the accessibility features for accessing the 3D world was considered easy to use (Sonar, GPS, Compass)

- The Backpack was easy to use. Since both the backpack and the GPS use a voiced menu system which also is used by other in game objects (like Communication devices) and also the game shell start menu, they are probably also easy to use.
- The game itself was perhaps too easy (question #7 and question #10, Blind 11), although the quests and enemies are considered "very challenging" (question #10 Blind,19).
- I have to follow up especially on user Blind,49 since he had not yet completed the game when responding.

4.2 Further work

Further work in this field is definitely needed. Hence I would like to say something about my thoughts for what future work need to be done.

More than 40 years has passed since the first computer games were developed (e.g. Spacewar from 1962). Yet you still have the same implied prerequisites to play a game, i.e. full sight, non-impaired hearing, cognitive and motor functions. In short the game industry excludes many (or most) disabled, potential gamers. Compared with the web industry (only about 10 years old), the game industry has done very little (if anything) in the accessibility area.

There are small companies and individuals who develop games for disabled. These games break new ground in interface design using unconventional means of input/output but they tend not to attract the mass market. The scope of these special games are a fraction of the scope of mainstream titles, where budget is perhaps the most important factor. All this has also been the case of Terraformers.

The problems of making games accessible are many. Research, development and governmental funding is needed to make it happen on a broader scale. One reason that make games especially hard to make accessible, is the fact that games use special development tools and APIs which don't work well, or at all, with standard accessibility tools like screen readers etc. Also many games are designed to challenge the gamer physically and/or mentally (with parameters like speed and precision, quests and puzzles), which runs counter to accessibility aims.

I feel confident that virtually any game can be made accessible for any disability. However, I also know it is a huge task for any game developer to make games (and especially real-time 3D games) accessible, and it is almost an utopia to make games accessible for literally everyone. By collaborating with the rest of the community I think it is possible to develop methods of making all game genres accessible for as many gamers as financially possible.

To do this, in May 2003 I took the initiative to the Game Accessibility Special Interest Group (GA-SIG) within the International Game Developers Association (IGDA, www.igda.com). In March 2004 I presented a draft white paper at two roundtable sessions at the Game Developers Conference (San José, USA, www.gdconf.com). Part of the white paper is based on a survey on the web, where I have collected and analyzed input from accessible game developers around the world. This includes games accessible for all disability groups. The white paper will be published in the summer of 2004, available for download at http://www.igda.org/accessibility/

Game accessibility is a new field for the game industry and real-time 3D games is among the most demanding computer environments to make accessible. Many other fields where real-time 3D is used will also benefit if we are able to develop general methods of making real-time 3D games accessible. This includes virtual reality applications, e-commerce, distance education etc.

My wish is to work together with the game/simulation/3D industry, but also with peers at universities to develop these methods of game accessibility. By presenting white papers like this and the IGDA GA-SIG white paper and participating at conferences like the GDC (www.gdconf.com) and ICDVRAT I hope to spread the word and find interested colleagues around the world.

Acknowledgements: Thanks to colleagues and students at the Multimedia programme, Department of Education, Stockholm University and of course to colleagues in the Terraformers project, now a company called Pin Interactive. Also thanks to Stig Becker at the Swedish Handicap Institute, the staff and judges at the Independent Games Festival and all people who have tested and given feedback, answered surveys and helped us in making this game.

5. REFERENCES

Gaver, W. W. (1994). Using and Creating Auditory Icons. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces* (pp. 417-446). Reading, USA: Addison-Wesley.

Mynatt, E. D. (1997). Transforming graphical interfaces into auditory interfaces for blind users. *Human-Computer Interaction, 12,* 7-45.

Winberg, F. & Hellström, S. O. (2000). Investigating Auditory Direct Manipulation: Sonifying the Towers of Hanoi. In *CHI 2000 Extended Abstracts*

Interactive Painting – an evolving study to facilitate reduced exclusion from classical music concerts for the deaf community

A Lewis-Brooks

Computer Science Department, Aalborg University Esbjerg,
Niels Bohrs vej 8, 6700 Esbjerg, DENMARK

tonybrooks@cs.aue.auc.dk

www.cs.aue.auc.dk

ABSTRACT

Exclusion from the joy of experiencing music, especially in concert venues, is especially applicable to those with an auditory impairment. There have been limited investigation into how to reduce the exclusion for this community in attending classical orchestra music concerts. Through utilizing computer technology and human machine interfaces (sensors and cameras) to stimulate complementary senses through interpretation it is possible to reduce this exclusion. Case studies are presented where the visual and tactile interpretation of the music is able to give new meaning and understanding for such people.

1. INTRODUCTION

Man, as yet, has no means of imagining the nature of his own experience in space. Until artists have provided him with adequate forms to express what he feels in space he will not know the meaning of experience. - Herbert Marshal McLuhan.

Music is obviously a hearing experience which means that those with a damaged auditive mechanism experience music differently than those with their sonic attributes intact. This often leads to their self imposed exclusion from the joy of attending a full symphony orchestra concert. Others, for example from within a special needs community, are also often excluded from such experiences. There has been very limited research into making classical music accessible to members of these communities. In an effort to study this exclusion problem the author initiated an inquiry following a workshop that he was asked to chair by the National Danish Television station and the European Film School. This inquiry involved a commissioned study from a New Zealand organisation based in Auckland who was interested in the Deaf community being given access to classical music. These two case studies are referenced in this paper with examples of a suggested solution to reduce this exclusion problem for classical music concerts.

This paper presents a further element, another piece of the jigsaw if you like, of the author's ongoing body of work titled SoundScapes (Brooks 2004a) which is overviewed briefly for the reader's orientation in section 2 - subsections of the various aspects of the work are also referenced under this section. The two case studies are presented in section 3 & 4 with interviews. Section 5 informs of the sensor technology used in the case studies. Due to the success of the overall study the concept and methodology was subsequently implemented into the therapeutic aspects of the research and this is briefly introduced in section 6.

2. SOUNDSCAPES

The view taken here is that research in the area of Neuromuscular control, biomechanical aspects of performance, the link between cognition and action, together with recent developments related to pathology and adaptation can inform rehabilitation methods through implementation of this concept. - Marshall, F.J. M.D. Neurologist

SoundScapes uses sensor and camera technologies so as to "activate" a volume of free air space where interference, in the form of movement by a human, creates events that result in changes to a multimedia environment. The human is positioned within the environment that is created from the use of a computer

 ISBN 07 049 11 44 2

workstation. The feedback is created to be "fun" for the human to interact with so as to attain an optimal level of immersed 'presence' from the *play scenario*. Philosophies from 'performance art' are exhibited alongside certain gaming psychologies within the concept and methodology and as such the research has often been referred to as a *research art form.*

The work is cored on human interaction and experts in the field have suggested potential closure of the human afferent efferent neural (sensorimotor) loop. The SoundScapes research is based upon a simple observable fact: Placing humans in an environment where they have "immediate" control (through movement) over sensual feedback of suitably interesting content (audible, visual, tactile…) is both aesthetically pleasing and capable of much therapeutic value that is 'evaluand' (a thing to be evaluated – the value to be determined (Lincoln & Guba, 1984)). The complementary observable fact is that nobody yet knows why in any individual case this should be so! (Beyond the anecdotal evidence that individuals given such control over their environment attain a level of self-motivation and coordination not expressed otherwise) – *Brain Empathy EU Expression of Interest*.

2.1 Performance Art

> *Performance Art is now at the forefront of current art. An astonishing increase in the number of works and venues around the world testifies to this art form as the chosen medium for articulating 'difference', whether dealing with issues of identity, multiculturalism or globalisation.* - Goldberg (2001)

'Performance' is an accepted genre within conceptual art that has a long history and embedded philosophies that acknowledges articulation of *difference*. Goldberg (2001) has documented extensively on the form of "Performance Art" showing that from the early Futurism movement at the start of the twentieth century to the present time the manifestos and philosophies which have been stated are centred on a freedom of expression that happens at a specific moment in time (real-time). The author implements such philosophies into the SoundScapes manifesto so as to further articulate and celebrate differences in humans through their empowerment in being able to create in real-time through the arts. This is achieved through interactive systems that are used for research in the human sciences and which are used in this study with classical music. The results give new opportunities within the special needs field giving additional voice and meaning to '*difference*' and the celebration of it. This fact is highlighted by the unsolicited communications written to the TV channels by viewers and audience as well as the study interviews with the audience, and performers, following the events.

2.2 Movement

Movement is an element of performance art and is an essential element of expression and joy which is often shown in response to music by humans. It is also an expression through movement that creates music, from the conductor's baton gesture, the bowing of a string, striking of a drum or even at its extreme in respect of a singer. This mapping between movement and music translates the expressive cues via a computer workstation such that they are capable of manipulating digital multimedia relative to the input from the human participant (performer) in *real-time*. The multimedia content can be tailored in a computer workstation so as to be relative to the context, the individual and the goal of the interaction.

2.3 The Question of Inter-relationship of content

Can we achieve a satisfactory relationship between correlations of music and image? - Is it essential for the interpretation of the experience of the music for the goals of the study? - If so what strategies should be targeted? - Will composers of the future create with equal weight across complementary sensory stimuli? -

Perplexity has a parallel to these and similar questions throughout the history of audiovisual research. It would seem that through the history of audiovisuals and the many people investigating the phenomena of sound and image that nobody has achieved a satisfactory outcome to date. A one to one relationship of mapping is quickly tiresome irrespective of methodology used to stimulate the senses. Kapuscinski (1997) has probably come closest to defining what is required when in "*Basic Theory of Intermedia - Composing with Sounds and Images*." he describes 'structurally integrated intermedia composition' where the unique property is the creative focus on the content emerging between the senses. In his PhD thesis he writes: 'This intermodal dimension is accomplished by linking, which means that while the freedom and individuality are the base for each dimension, it is the relationships between dimensions that constitute the medium.' Related to an Kapuscinski's intermedia theory is an article titled "*Sense and Intersensorality*" for the Leonardo publication *Intersenses and New Technologies*, where François Delalande discusses the more general problem of musical meaning where he asks as to how sound relates to something outside of the world of sound. He relates to temporal form such as that of movement and the relationship to sound 'forms' (shapes, profiles) and this is clearly related to the earlier research involving the author (Brooks *et al* 2002) which featured movement relative to sound 'forms' that were used in therapy. Jean-Pierre Ternaux who discusses in the same journal the unusual phenomenon of "*Synesthesia: A Multimodal Combination of the Senses"* - where he investigates the pathological context of synesthesia as well as relating to it as a physiological behaviour that involves the multimodal combination of the senses and in so doing approaches it from the cerebral mechanisms angle. Also in the same issue Rolf Inge Godøy in his article titled "*Motor + Mimetic Music Cognition"* highlights the well known fact that there is more than just the auditory aspects that enrich the experience of music stating that "... in particular, images of movement appear to be deeply embedded in our perceptual and cognition of music. Explorations of mental images of music related movement could enhance our understanding of music as a phenomenon, as well as be of practical value in various music-making tasks." In this study however, while being aware of potentials from such correspondences and relationships across the senses the strategy was initiated to resource the musical scores, audio recordings of the work (if available) and words, where appropriate, and use this material for the creation of various interpretable digital image libraries for real-time improvisation by the author.

3. DENMARK

The workshop in Denmark was as a consequence of a survey with members of the public who had an interest in classical music. The results from the survey showed that audiences watching classical music on television were declining in number. The workshop gave insight into potential ways to inspire the public to return as viewers and ways to increase the audience. The workshop study was hosted at the European Film School by the National TV station's classical music production department.

3.1 The TV studio study

A professional choir was selected for the TV studio. The music composition was supplied with text and a recording from the conductor. The resulting visual sequence was a series of computer graphics depicting a gothic cathedral window with coloured leaded glass. These colours could be morphed by the conductor's hand gesture. The choir consisted of around forty singers of mixed gender and aged between 20 and 50 years. Instructions were given to the conductor that the experiment was with an invisible interface to a computer and that as he had forty singers relying on his conducting there should be no conscious effort to change any of his gesture due to the image manipulation.

3.1.1 Technique and method. Ultrasonic (Soundbeam®) sensors, a volumetric 3D infrared light sensor unit (author prototype) and cameras were used. The mapping of the data from the hand movement was programmed to open filters Red, Green & Blue in the computer image sequence. A gesture between sensor spaces resulted in proportional opening of the filters and a resulting blend of the colours. No movement within the active sensor space from the conductors results in black as no filters are opened.

Figure 1. *Denmark classical choir – conductor (left - raised) gestures and digitally paints the background. The choir members (right) follow the gesture as directing their singing.*

The conductor was set up on a raised podium in front of where the choir was positioned in front of a 12m wide x 8m high screen. The conductor was asked to wear a white paper painting suit to maximize his amalgamation into the projected images. The choir (who arrived for the TV shoot with tuxedos and black formal dresses for the TV broadcast!) was asked to wear white sheets with holes for their head so as to maximize projected light reflection. Two LCD (Liquid Crystal Display) projectors were set up on the main floor space at approximately sixty degrees angle behind the conductor and angled at around twenty degrees from the horizontal. The projected image was thus raised above the heads of the choir so as to have a clear projection area, and to minimise projected light in their eyes. The two main projected images were vertically adjoined so as to extend the effect and to give the effect of one projection. This use of two projectors also gave two shadow images of the conductor at the centre of each of the projections, this was to give viewers the opportunity to observe the hand gestures even if the camera angle was behind the conductor and thus eliminating a view of his hand gesture. The conductor was advised to make a purposeful and defined entry into the active sensor space so as to give a dynamic opportunity for viewers to correlate the causality of the interaction. Similarly at the close of the piece he was advised to slowly remove his hands from the sensor space so that the last tones heard from the choir ceased simultaneous to a fade to black of image.

3.1.2 Result. .From observing the conductor during the experiment it became obvious that his gestures were in fact becoming more emphatic, and animated. Interviews with choir members confirmed this but they agreed it did not detract from his guidance. Each subsequent reiteration of the piece of music was progressively initiated with a more vibrant entrance into the sensor space by the conductor. When interviewed after the sessions were completed the comments from the conductor stated that it felt "… *like a power trip,*" and "… *as if the air was electrically alive in front of my body*."

4. NEW ZEALAND (NZ)

The study in New Zealand was with a symphony orchestra based in Auckland. The series of four concerts were titled "Four Senses" and had a defined target group of the Deaf community with the goals of giving opportunity for this group to have an improved experience of classical music. In addition to the orchestra there was a mixed ability contact-improvisation dance company, a singer (visually impaired), a signing choir (for interpreting the singer's text via New Zealand sign language). There were also vibration cushions in the first four rows in the auditorium that received the music from the orchestra via the sound mixer station. Balloons were also given to selected audience with impairment so that they could lightly hold and feel the vibrations from the music. The author was invited to participate by artist and lighting director Raewyn Turner who implemented her methodology of interpretation of music into light and olfactory stimulation via custom aerosol input to the air conditioned system of the auditorium. Publication of her methodology including use in this series of "Four Senses" concerts is documented (Turner 2003). The concerts are documented further in Brooks (2004a) and Brooks and Turner (2003).

4.1 Technique and method

The set up is described in the publications mentioned in the previous paragraph but suffice to say that the expressive information from the performers were collected via the cameras and sensors and manipulated by the author into a real-time free interpretation of the music juxtaposed to the lighting changes. In the Denmark study only the conductor was involved in the digital painting. In New Zealand it was also members of the orchestra, the dancers, the singer and the signers that were involved. The goal was to show how the emotive expression that occurs as a result of performing music could be translated through interpretation for the audience in a large 'live' environment.

4.1.1 Result. The conductor in Auckland was observed to again very quickly be at ease with the system and there was no interference reported, however the use of sheets of music notation (turning over) in certain pieces of music disrupted the sensors continuous data stream capture of expression. Initially it was thought to distract from the objective of the exercise, however it did in fact become an integrated aspect, as that was also his expression at that moment relative to the interaction to the orchestra.

The interviews of some of the audience were included in the National TV broadcast. A number of letters was sent to the TV station (three examples below), which was an unexpected addition to the interviews. The orchestra performers on stage stated that they would have liked to be in the audience to experience the event. One musician in particular who herself had a hearing impairment expressed this very strongly and that all she had spoken to in the community had only positive remarks.

Figure 2. *The New Zealand mixed ability dancers (left) and orchestra members (right) – all interactively digitally painting the backdrop from their expressive gesture.*

Three examples of the mails received (published with permission – addresses deleted):-

----

I accompanied five people with varied degrees of deaf blindness and their caregivers to the Friday performance of the 'Four Senses" concerts. I was impressed, moved and delighted at everything that I with my senses intact and unimpaired could experience. More importantly, it was so rewarding to observe how much of the music, lights, smells and movement our disabled friends were able to enjoy in the extraordinary ways your team made available. At a meeting of parents this week we heard from the caregivers of the gratifying response that their very handicapped charges made to this fantastic sensory experience. Delight was the common reaction, it seems, and also a stillness that they rarely show that indicated they were receiving the stimuli. I have been asked to express the heartfelt thanks of all members of the Quality of Life Trust for your gift to the sensory-impaired people of Auckland. May you be able to repeat the concerts very soon!

Today I was deeply moved watching the enjoyment that the audience experienced in an outstanding and innovative concert. The tears rolled down my face as I watched the expression on the faces of those who participated. Not only was the concert beneficial to those with disabilities, but it was also inspiring to people such as myself who have an empathy for those less able. Please give me the contact names and addresses of those involved in the technical & visual production as I would like to organised a similar concert for people in the Wanganui, Palmerston North & New Plymouth areas. Armed with this information I will challenge local Rotarians to put on a charitable concert that will involve school orchestra's and other community organisations. I look forward to receiving any information you can provide. This will be the first step in realising a vision. Thank you for screening such a wonderful show.

To whom ever this may concern I watched your programme which aired 16/06/02 on Sunday, it was great and I was wondering if I could get further information on who was involved in putting a visual and feeling concert together, as well as info on how they did it, and in particular the aids which were used by the audience, i.e. vibrating cushions. I believe that, a show such as this would be great in the region of Southland, where I come from, although, being so far from the main centres, persons with disabilities can become isolated from such beneficial and fun activities.

The response in the many interviews from both the Danish and New Zealand study gives credence to the concept of utilizing interactive complementary sensory stimuli in live performance to give accessibility to audiences that may otherwise be excluded. The Interviews point to the fact that multi-sensory stimulation does enhance the experience for not only the impaired community but also, as can be read in the mails included in this paper, for others. Further consequences are suggested as more composers in the twenty first century are becoming more aware of the studies that are underway exploring such enquiry and they begin composing specifically to account for mixed phenomena opportunities. Such explorations in composing could result in their writing with a specific inclusion of sensory stimulus beyond solely sonic. Scriabin with his synesthesia related composition *Prométhée, le Poème du Feu* is an example of such earlier explorations by composers. As a result of the case studies which was built on earlier work with multi-sensory multimedia interaction in therapy (e.g. Brooks *et al* 2002) the author suggests this study further adds to the field of rehabilitation and habilitation where environments that are pleasing to the patient are able to be successfully controlled by the patient. This success is a factor of the immersive ‘presence’ and alongside the embedded layers inherent to the system where various elements can be quantified a platform for future collaborative research is suggested. The quote from Marshall above, who is a leader at one of the top movement disorder Centres in the World (Rochester University Hospital) highlights the interest from the medical side.

5. SENSOR DETAILS

The Soundbeam® (below left) is a mature and sophisticated ultrasound movement sensor that is detecting movement along its linear axis. The system enables accurate reproduction of desired events. It translates movement to MIDI and is capable of responding to any MIDI control message and program change. This enables remote change of sensitivity which is a main factor when working with expressive gesture interpretations. The Soundbeam® is stable under varying light conditions and in a theatre auditorium with light shows this is important (as for these case studies). There is a minor problem with an auditive buzz which is common with such devices but this did not interfere in any way from its use.

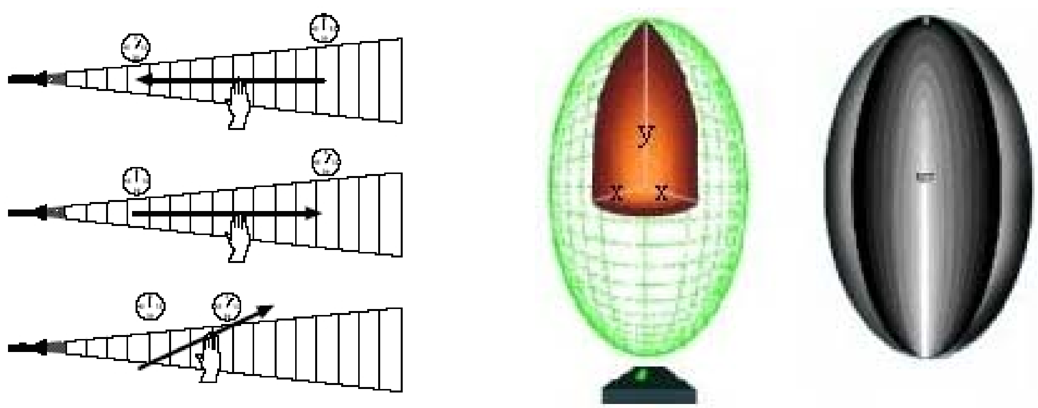


Figure 3. *(Left) Soundbeam® sensor where various gesture along linear axis articulates various predictable feedback. (Right) 3D infrared sensor ‘active’ space.*

The infrared sensors (above right) utilized in the studies are volumetric 3 dimensional (3D) and like the Soundbeam® use the MIDI protocol. The maximum resolution of the sensor is 0 – 127 along its central “core” axis (Y) and 0 – 127 from the centre axis to the “skin” (X). The depiction in the figures is to illustrate the shape and maximum resolution of the active space. The beam is sensitive to lighting change which was a problem in the auditorium where light changes were inherent to the production. If a specific event is desired then one must traverse through the outer skin to the point in space where the desired note resides – this is not ideal. Programming can alleviate this. Further detailed in Brooks (2004e).

6. THERAPY

As a result of the two case studies the methodology was adapted for various sessions within the disabled community. The digital painting was supplemented by control of intelligent robotic devices (Brooks 2004d) which proved a powerful physical feedback for motivated interaction which can be seen in figure 4. It was also adapted into cross modal painting as detailed in Figure 5.

Interactive digital technologies which capture movement and subsequently control audio visuals from the interface have been investigated primarily for dance, performance and music with the result from the research of a number of cross platform programs available for consumer exploration (Appendix 1). This gives opportunities for parents and interested therapists to create specific tools for their needs.

The “Interactive Painter” concept and method is presented in “*Inspiration Day*” therapist training seminars in Denmark as well as in the author’s SoundScapes global presentations and workshops. Brooks (2004a) informs of a similar set up that was used in an art performance in New York. The initial therapeutic use is detailed in Brooks (2004b). Custom systems and training are available from the author on request. Research collaborations are welcome as are presentation and workshop invitations.

Figure 4. *(Left) a session where father and brother look on as the girl makes music and paints. (Right) The boy watches his digital painting and robot control with just his hand (sensor assembly above).*

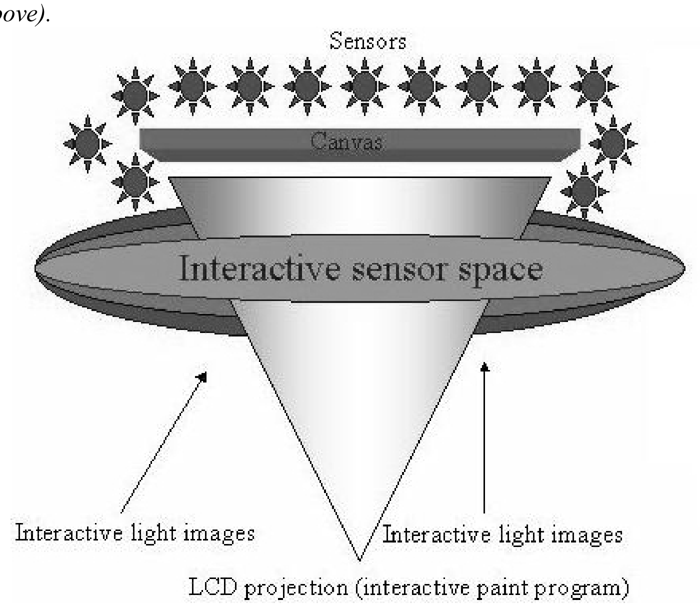


Figure 5. *Adapted to Cross Modal Painting where traditional water colours are used with canvas and brush in conjunction with the digital "Interactive Painter" for therapy within the disabled community. Movement with a paint brush is detected by the sensors as it enters the proximity of the canvas. This triggers dynamic multimedia that is affecting and motivating for the user. An interactive causal loop is thus established where proprioceptive sense (body awareness) is amplified so as to aid in therapy.*

7. CONCLUSION

To summarize the contribution of this paper it is left to an episode extracted from the TV production from New Zealand. As mentioned in the text a hearing impaired member of the orchestra from the New Zealand event is interviewed in the TV production. The conductor (and musical director) is also interviewed and purposefully enlightens the TV audience to an episode when the same member of the orchestra exits the stage in tears following one performance. When confronted about the problem by the worried conductor the young musician says that "*You don't know what this means to people like us*" – the conductor concludes his interview emotionally by stating – "*I think that says it all*"!

Acknowledgements: The author wishes to thank all involved in the Danish and New Zealand productions. Thanks also to the public cited (and not cited) in the text who wrote about their experiences - and those involved in the taped interviews – *Sorry but I could not include them all in this paper*!

8. REFERENCES

A L Brooks (1999) Virtual Interactive Space (V.I.S.) as a Movement Capture Interface Tool Giving Multimedia Feedback for Treatment and Analysis, *Proc. 13th Int. Congress of World Confederation for Physical Therapy*, Yokohama, Japan, pp. 66.

A L Brooks (2002) 'Soundscapes' as a Human Computer Interface for Use in Therapy, Elderly Care and Rehabilitation, *Proc. 5th Asia Pacific Conference On Computer Human Interaction, User Interaction Technology in the 21st Century,* 2002, Beijing, China, pp 833-844.

A L Brooks, S Hasselblad, A Camurri, N Canagarajah (2002) Interaction with shapes and sounds as a therapy for special needs and rehabilitation, *Proc. 4th Int. Conference On Disability, Virtual Reality, and Associated Technologies,* Veszprém, Hungary, pp. 205-212.

A L Brooks and R Turner (2003) A Collaboration, *Intl. J. Wegway, Primary Culture* **5**, www.wegway.com.

A L Brooks (2004a) Soundscapes, In *Inhabited Information Spaces – Living with Your Data,* D N Snowdon,

E F Churchill and E Frécon (Eds.), Springer-Verlag London, pp. 89-100.

A L Brooks (2004b) CAREHERE - Creating Aesthetically Resonant Environments for the Handicapped, Elderly and Rehabilitation: Sweden, *Proc. 5th Int. Conf. On Disability, Virtual Reality, and Associated Technologies,* Oxford, England. *In press*.

A L Brooks (2004c) HUMANICS 1: A study to create a home based telehealth product to supplement acquired brain injury therapy, *Proc. 5th Int. Conf. On Disability, Virtual Reality, and Associated Technologies,* Oxford, England, *In press*.

A L Brooks (2004d) Robotic synchronized to human gesture as a virtual coach in (re)habilitation therapy, *Proc.Int. Workshop for Virtual Rehabilitation,* Lausanne, Switzerland, *In press*.

A L Brooks (2004e) Enhanced Gesture Capture in Virtual Interactive Space, *Proc. Computers in Art, Design and Education,* Copenhagen, Denmark, http://asp.cbs.dk/cade2004.

S Brown, B Merker and N L Wallin (1999) *The Origins of Music*, MIT Press, pp. 3-6.

F Delalande (2003) Sense and Intersensorality, *Intl. J. Leonardo*, MIT Press, **36**, 4, pp. 313-316.

R I Godøy (2003) Motor + Mimetic Music Cognition*, Intl. J. Leonardo*, MIT Press, **36**, 4, pp. 317-319.

R Goldberg (2001) *Performance Art – From Futurism to the Present*, Thames Hudson, Back cover.

Y Lincoln and E G Guba (1984) *Naturalistic Inquiry*, Sage publications, p 227.

J Kapuscinski (1997) Basic Theory of Intermedia - Composing with Sounds and Images, *Monochord, De musica acta, studia et commentariti*, Marszelek Pubs., **XIX**, pp. 43-50.

J P Ternaux (2003) Synesthesia: A Multimodal Combination of the Senses, *Intl. J. Leonardo*, MIT Press, **36**, 4, pp. 321-322.

R Turner (2003) Olfactory Translations and Interpretations, *Int. J. Performance Research*, **8**, 3, pp 104 –112.

APPENDIX

Dance and Music research technologists and developers with commercial audiovisual interactive programs:

- Camurri, A "Eyesweb" http://www.eyesweb.org
- Siegel, W. "DIEM Digital dance System" http://diem.dk/english/diem/digdance/info.php
- Rokeby, D. "Very Nervous System" http://homepage.mac.com/davidrokeby/articles.html
- Coniglio, M. "MidiDancer" http://troikaranch.org/technology2.html
- Weis, F. "Eyecon" http://www.eyecon.de

Disclaimer Statement: As stated in the text this paper is a further piece of the evolving SoundScapes research jigsaw. In presenting new and unpublished aspects of the concept and methodology it cross-references the author's prior art which is detailed in the referenced publications list under his name. Such cross-reference is deemed unavoidable and in an attempt so as to minimise duplicate publication the reader may experience brevity of various sections in this paper. Hopefully this does not deter from the flow of the piece and the understanding of the content. TB.

ICDVRAT 2004

Session IV. Enhancing Mobility and Accessibility

Chair: Tamar Weiss

Simulating the effects of fatigue and pathologies in gait

T Komura[1], A Nagano[2], H Leung[3] and Y Shinagawa[4]

[1,3]Department of Computer Engineering and Information Technology,
City University of Hong Kong, 83 Tat Chee Ave, Kowloon, HONG KONG

[2]Computational Division, RIKEN,
2-1 Hirosawa Wako, Saitama, JAPAN

[4]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign,
405 N. Mathews Avenue, Urbana, Illinois 61801, USA

[1]*taku@ieee.org*, [2]*a-nagano@riken.jp*, [3]*howard@cityu.edu.hk*, [4]*sinagawa@uiuc.edu*

[1]*www.it.cityu.edu.hk/~itaku*, [2]*www.it.cityu.edu.hk/~howard*, [4]*netfiles.uiuc.edu/sinagawa/www/*

ABSTRACT

In this study, we propose a new method to simulate the effect of fatigue and pathologies in human gait motion. The method is based on Angular Momentum inducing inverted Pendulum Mode (AMPM), which is the enhanced version of 3D linear inverted pendulum mode that is used in robotics to generate biped locomotion. By importing the gait motion captured using a motion-capture device, the value of AMPM-parameters that define the trajectory of the center of mass and the angular momentum are calculated. By minimizing an objective function that takes into account the fatigue and disabilities of muscles, the original motion is converted to a new motion. Since the number of parameters to define the motion is small in our method, the optimization process converges much more quickly than in previous methods.

1. INTRODUCTION

Computer simulations based on musculoskeletal models are often used to analyze the role of specific muscles during motions. Many researchers have analyzed the contribution of some specific muscles to the performance of certain motions by using optimization methods based on forward dynamics. For example, Pandy and Zajac (1991) analyzed the role of biarticular muscles in maximal jumping, and they concluded that biarticular muscles contribute to jumping performance by redistributing segmental energy within the musculoskeletal system but do not contribute to the energy of jumping. Neptune et al. (2001) analyzed the role of the plantarflexor muscles during gait, and they calculated the degree to which these muscles contribute to propelling the trunk in the forward direction. Human gait motions have also been simulated by combining forward dynamics and optimization (Yamaguchi and Zajac, 1990; Anderson and Pandy, 2001; Neptune et al., 2001). Although these researches have succeeded in simulating realistic normal gait, there are several shortcomings with these methods that make it difficult to apply them for the diagnosis of patients due to the following reasons:

1. First of all, researchers must provide a good initial guess of the activation data to prevent the optimization process from getting caught into some local minima. In most of the cases, this is done through trial-and-error by researchers. This is quite time consuming, and the researcher needs special experience to determine the good set of muscle activations to generate realistic gait motion. In addition, since this process is done completely manually, the motion can become subjective to the researcher.
2. Next, since all the time sequence data for the muscles must be determined through the optimization, costly computation is required until the result is converged to the optimal value. Neptune (1999) reported that more than thousands of iterations are needed before the optimized value is obtained.
3. Third, although balance-keeping is one of the most important factors for gait motion, there was no explicit way to describe the balanced motion in these researches. As a result, the search space for the parameter is much larger since it also includes those parameter values that lead to unbalanced motion.

In robotics, biped locomotion of humanoid robots is one of the most exciting topics these days. Many researchers have developed humanoid robots that are capable of performing biped gait motion (Kajita et al., 1992, 2001). A method commonly used to generate gait motion is the 3D Linear Inverted Pendulum Mode (3DLIPM) (Kajita et al., 1992, 2001). The great advantage of the 3DLIPM approach is that the trajectory of the center of mass (COM) can be written in an explicit form. The 3DLIPM approach is a hierarchical approach, in which the abstract motion such as the trajectory of COM is determined first, and the details of the motion such as the kinematic data are calculated later. The advantage of the hierarchical approach is that the controller does not have to waste time determining low level control signals such as the torque exerted at the joints or the muscle activation data. With the hierarchical approach, it is possible to simplify complex models of humanoid robots that have too many degrees of freedom (DOF) to be controlled directly.

In addition, as the motion of the COM is explicitly controlled, the balance of the humanoid robot is assured in the feedforward stage. However, no angular momentum has been generated by the 3DLIPM since it assumes that the COM is a mass point and that the ground force vector always passes through the COM of the system (Figure 1(a)).

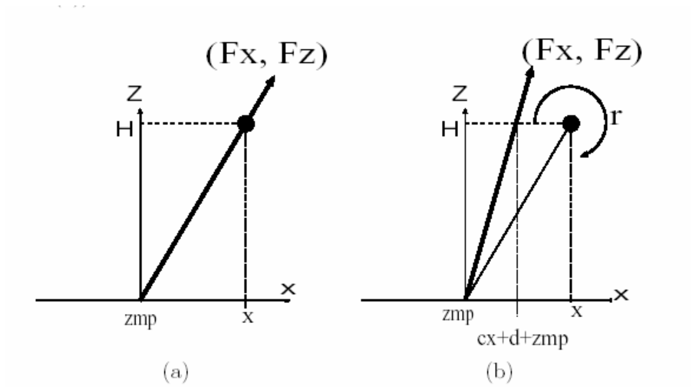


Figure 1. *The standard 3DLIPM (a) and AMPM (b). The 3DLIPM restricted the ground force to pass through COM while the AMPM allows the horizontal component of the ground force to have linear relationship with the position of the COM.*

It is known that angular momentum is consistently generated around the COM when humans walk. The amount of angular momentum gets even larger as the motion involves larger bending of the thorax, which can be often observed in pathological gait. It is therefore difficult to model human gait motion by using 3DLIPM. To represent trajectory of COM and angular momentum of human gait motion in an explicit form, we have extended the 3DLIPM model so that angular momentum can be induced by the ground reaction force (Kudoh and Komura, 2003). This model is called *Angular Momentum inducing inverted Pendulum Model* (AMPM). In this study, we propose a new hierarchical approach to calculate human gait motion using AMPM. By combining AMPM with the musculoskeletal model, it is even possible to calculate pathological gait. The effects of deactivating the gluteus medialis muscles were simulated and compared with the motion of patients with similar disabilities.

The trajectory of the COM and angular momentum is first calculated by using AMPM. Next, the trajectories of the generalized coordinates including the position of the pelvis and the joint angles of the whole body are determined using inverse kinematics. After determining the muscles of the body to be deactivated, an objective function based on the time series data of the force exerted by this muscle is formed. By searching for an optimal set of AMPM parameters that minimizes the objective function defining the gait motion, the pathological gait motion is calculated.

The method proposed in this study has the following advantages compared with previous methods:

- Since the gait motion is described by the AMPM parameters, there is no need to specify all the input parameters of the muscles. As a result, the computational cost for the optimization is much less than previous methods.
- As the balance of the human body model is explicitly kept by using the AMPM model, the optimizer only needs to search for the optimal set of parameters in terms of muscle force, and hence does not need to go through a large number of trials to generate the balanced motion.

2. METHODS

In this section, the method to simulate the effects of pathologies and fatigue during human gait is explained. First we review the Angular Momentum inducing inverted Pendulum Model (AMPM) (Kudoh and Komura, 2003) and also explain further enhancements so that it can handle imported human motion data. Next, the details of the musculoskeletal model we used in this research are explained. Finally, the method to simulate the effects of pathologies and fatigue by optimization is explained.

2.1 Angular Momentum inducing inverted Pendulum Model

The AMPM enhances the 3DLIPM in a way the ground force vector is calculated to be not only parallel to the vector connecting the ZMP and the COM; its horizontal element can be linearly correlated to the ZMP-COM vector (Figure 1(b)). As a result, rotational moment will be generated by the ground force. In the previous studies, we assumed the ground force vector to be linearly correlated to the position of the COM:

$$\ddot{x} = Ax + B \tag{1}$$

As we assume the single support phase of human gait follows this rule, the parameters *A* and *B* can be calculated using the position and acceleration of the COM at the lift-off and heel-strike during gait. However, when actually we substitute data into this equation, the trajectory of the acceleration does not match well with those by humans. As shown in Figure 2, oscillatory gap is observed between the acceleration of the COM and that generated by AMPM.

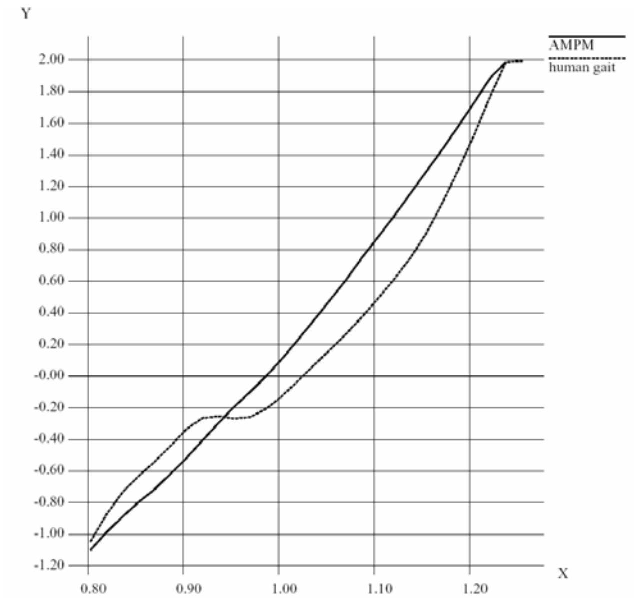


Figure 2. *Acceleration of the COM during gait (dashed line) and that by AMPM (bold line).*

In order to decrease the gap between the trajectory by the AMPM and the trajectory of the COM during gait, we propose to add trigonometric functions in the right side of Equation (1) as follows:

$$\ddot{x} = Ax + B + C\cos\omega t + D\sin\omega t \tag{2}$$

The parameters *C, D, ω* are decided in a way that trajectory of the differential equation matches well with the trajectory of the COM. Equation (2) has the explicit solution as follows:

$$x = -\frac{B}{A} + C_1 e^{-\sqrt{A}t} + C_2 e^{\sqrt{A}t} - \frac{\cos\ \omega t + D\ \sin\ \omega t}{A + \omega^2} \tag{3}$$

where C_1, C_2 are constant values which can be calculated from the terminal conditions. In this study, we make the following two assumptions through observation of human motion:

- The position of the ZMP is linearly correlated to the position of the COM. Therefore, the position of the ZMP can be defined by (*zmp,* 0) = (*cx* + *d,* 0), where (*c, d*) are constant parameters.

- The height of the COM can be expressed by the following equation:

$$z_g = H + z_h \sin\frac{t}{T}$$

where H is a constant value, T is a half cycle of gait motion, z_h is the amplitude of the motion, and z_g is the height of the COM.

Then, the ground force vector can be written as

$$F_x = m\ddot{x} = m\left(AC_1 e^{-\sqrt{A}t} + AC_2 e^{\sqrt{A}t} + \frac{\omega^2}{A+\omega^2}(\cos\omega t + D\sin\omega t) \right)$$

$$F_y = m\left(g - \frac{1}{T^2}\sin\frac{1}{T} \right)$$

where m is the mass of the system. The rotational moment r around the y-axis can be calculated by

$$r = ((1-c)x + d)F_y - (H + z_h \sin\frac{t}{T})F_x$$

Using this moment, it is also possible to calculate the angular momentum generated by the rotational momentum in an explicit form, although we cannot list it here due to its length. This form was calculated using the software *Mathematica*. AMPM is used to represent the motion of the COM/ZMP/angular momentum in the frontal and sagittal plane. Then, the generalized coordinates of the body are then calculated using inverse kinematics (Kudoh and Komura, 2003).

2.2 Musculoskeletal Model

To simulate the motion by pathological patients, it is necessary to prepare a physiological model of the human body. The musculoskeletal model developed by Delp (1990) was used in this study. This data includes the attachment sites of 43 muscles on each leg and physiological parameters such as the tendon slack length, optimal fiber length, pennation angles, and maximum exertable force. Muscles are attached only to the legs, and no muscles are put on the upper half of the body. The upper half of the body is composed of the thorax, head, upper arms, lower arms, and hands.

However, only the thorax (3 degrees of freedom (DOF)) and the upper arms (3 DOF each) are allowed to move among these segments. The lower half of the body is composed of the pelvis, and the femur, tibia, patella, talus, calcaneous, and toes in each leg. The joints of each leg are composed of a 3-DOF gimbal joint (hip joint) and 1-DOF joint (knee, ankle, calcaneous, and metatarsal joint). Therefore, the total degrees of freedom of the body, including the 6DOF of the pelvis in the global coordinate system, is twenty nine.

The musculotendon model of Hill (1938) is used, and parameter values were derived according to Delp (1990). The musculotendon model is composed of three elements: a contractile element (CE, representing all the muscle fibers), a parallel elastic element (PEE, representing all connective tissues around the muscles fibers), and a series elastic element (SEE, representing all series elasticity, including tendons). At each time step, the musculotendon length was determined from the posture (i.e., as a function of joint angles). Thereafter, it is possible to calculate the maxima and minima that bound the muscle force at each time step:

$$F_m^{\min} \le F_m(t) \le F_m^{\max} \tag{4}$$

where $F_m(t)$ is the musculotendon force of the m-th muscle, $F_m^{\min}$ and $F_m^{\max}$ are the minimum and maximum force developed by this muscle, which are determined by their force-length-velocity properties:

$$F_m^{\min} = f(F_m^0, l_m, v_m, 0)$$
$$F_m^{\max} = f(F_m^0, l_m, v_m, 1)$$

where function $f(F_m^0, l_m, v_m, a_m)$ is the force-length-velocity surface assumed in the musculoskeletal model (Zajac, 1989), F_m^0 is the maximum force constant calculated from the average cross-sectional area of the muscle, l_m is the length and v_m is the velocity of the m-th muscle being contracted, and last parameter (0 and

1) in these equations is the activation level of the muscle a_m that determines the amount of force exerted by the contractile element ($0 \le a_m \le 1$).

2.3 Calculating the effects of fatigue and pathologies

Torque $\tau_j(t)$ developed at joint j is theoretically generated as follows by the muscles crossing the joint:

$$\tau_j(t) = \sum_m F_m(t) r_{m,j},\ j = 0,\ldots,n_{dof} \tag{5}$$

where $r_{m,j}$ is the moment arm of muscle m about the j-th joint axis, and n_{dof} is the number of degrees of freedom whose torque are assumed here to be generated by the muscles. They include the flexion/extension, adduction/abduction, and rotation at the hip, flexion/extension at the knee, and plantarflexion/dorsiflexion at the ankle, and therefore, by taking both legs into account, $n_{dof} = 10$.

The muscle forces at each time step can be calculated by minimizing the muscle stress (Crowninshield and Brand, 1981):

$$J = \sum_m^{n_m} \left(\frac{F_m(t_i)}{F_m^0} \right)^2 \tag{6}$$

where n_m is the total number of muscles ($n_m = 43 \times 2 = 86$). J was minimized using quadratic programming, which is a method to minimize a quadratic form while satisfying linear equality and inequality constraints. In summary, the muscle forces can be calculated by minimizing Equation (6) while using Equations (5) and (4) as constraints.

To describe the process to calculate the pathological gait, a new function is defined here that summarize all the processes explained previously, including the AMPM, inverse kinematics, inverse dynamics and static optimization:

$$F_m = f_m(p,t)(m = 0,\ldots,n_m - 1) \tag{7}$$

where $\boldsymbol{p}$ is the vector composed of parameters defining the gait motion by the AMPM, and F_m is the force by muscle m calculated by minimizing Equation (6). To simulate the effect of weakening muscle m, the following criterion is minimized until it is smaller than the threshold value:

$$J_m = \int_0^T \left[f_m(p,t) + \alpha(p,t) \right] dt \tag{8}$$

where T is one cycle of the gait motion, and $\alpha(\boldsymbol{p}, t)$ is a penalty function based on the external torque that has to be applied to the body to assist the musculoskeletal model accomplish the motion when there is no solution found for the static optimization problem at time t (Komura et al, 2000; Komura and Shinagawa, 2001). The penalty function helps to avoid the motion to converge to one that is not feasible by the musculoskeletal model. This optimization is done by sequential quadratic programming.

3. EXPERIMENTAL DATA ANALYSIS

To examine the validity of the method proposed in this paper, we have generated normal and pathological human gait motion using our method, and compared the kinematical and dynamical data with those by humans.

First of all, the normal gait motion that was captured using a VICON motion capture system was imported, and the AMPM parameters were calculated. Next, by minimizing an objective function based on the gluteus medialis that has the form of Equation (8), the effects of weakening the gluteus medialis during gait were simulated. As the optimization proceeds, features known as lateral trunk bending appears in the motion. The trunk swings from one side to the other, producing a gait pattern known as waddling. During the double support phase, the trunk is generally upright, but as soon as the single support phase begins, the trunk leans over the support leg, returning to the upright attitude again at the beginning of the next double support phase.

The trajectory of the gait motion before and after the optimization is shown in Figure 3 (a) and (b) as well. The objective function converges to the optimal value within one hundred iterations.

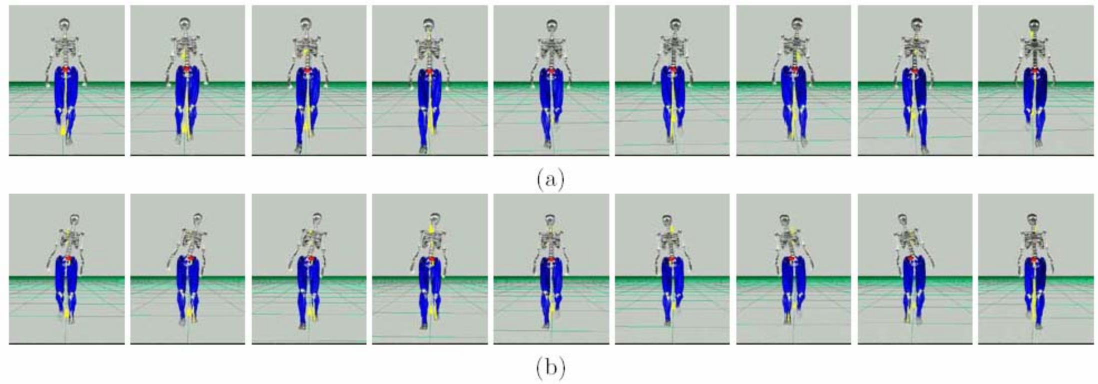


Figure 3. *The trajectory of the AMPM-generated motion before (a) and after (b) optimization.*

Next, in order to simulate the effect of fatigue, the maximum amount of force exertable by the muscles that are used during gait were gradually decreased. As a result, the strides of the gait gradually decreased as shown in Figure 4.

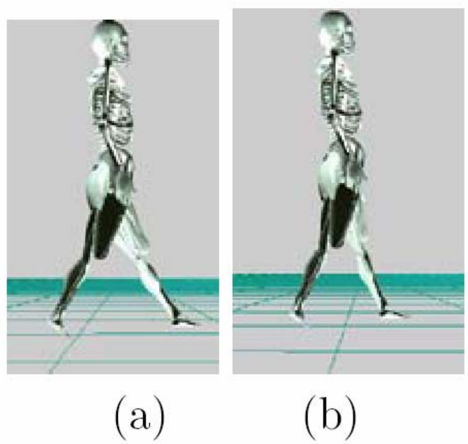


Figure 4: *Simulating the effect of fatigue during gait. The strides before the optimization (a) is shortened after reducing the upper bound of the muscle force that is used during gait (b).*

4. DISCUSSIONS

A method to calculate pathological gait by combining the musculoskeletal system and the inverted pendulum mode was proposed in this research. The motion is planned by the AMPM, and it is evaluated by using the musculoskeletal model. As the number of the parameters to determine the motion is less than previous methods, the optimization process converges more quickly than previous methods.

In our experiment, by minimizing the amount of force exerted by the gluteus medialis during the gait motion, the waddling gait motion was automatically induced. This is a natural phenomenon, as it is known that weak abductor muscles cause waddling. Waddling is a well-known abnormality of gait motion which is caused not only by weak abductor muscles, but also by congenital dislocation of hip joints and pain in the joints. Waddling reduces the torque and the bone-on-bone contact force at the hip of the support leg during the single support phase. The muscle force history of the musculoskeletal model performing normal gait and waddling gait calculated using static optimization are shown in Figure 5.

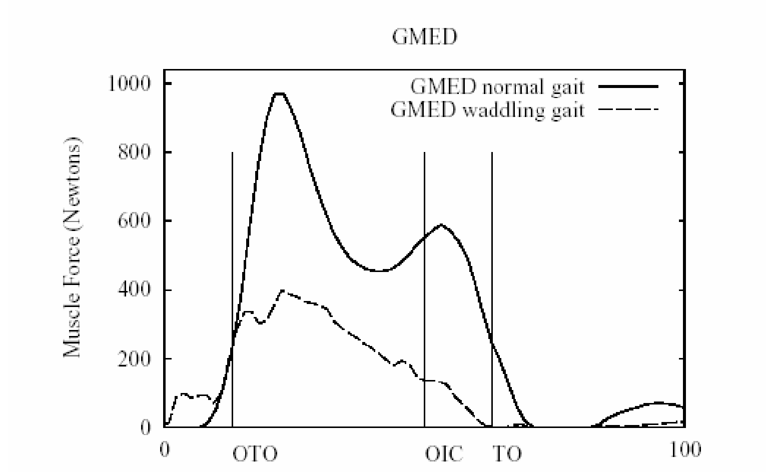


Figure 5： *The gluteus medialis (GMED) force calculated by static optimization method using human motion by a healthy subject (bold line) and by a patient with congenital dislocation (dashed line).*

Compared with dynamic optimization methods that have recently been used in biomechanics to simulate human gait and estimate muscle force, the method proposed in this study has the following advantages.

1. Our method to calculate human gait motion requires only a small number of parameters that define the motion of the AMPM. For example, if the musculoskeletal system is composed of 86 muscles, and during the gait cycle each muscle is allowed to change its input value 100 times, the total number of parameters that define the whole motion becomes 8,600. This means at least this number of iterations are needed for calculating the derivative of the criteria of optimization. Although techniques such as using the same input signals for muscles with similar roles (Pandy et al., 1990), or controlling muscles by “Bang-Bang” methods to reduce the search space (Yamaguchi and Zajac, 1990), are used, the number of parameters still remains very large. On the other hand, when using the proposed method, the number of parameters is less than ten. Since the number of the parameters is small, the optimization converges much more quickly than dynamic optimization methods. Comparing the amount of computation needed, our method requires less than one tenth of the forward dynamics approach.
2. Since the balance of the motion is assured by the algorithm of the motion generation, the optimizer does not have to spend further effort in keeping the balance through the gait cycle. This is one of the most important contributions of the paper, because when using optimization methods based on forward dynamics to simulate bipedal gait, most of the iterations are done for unbalanced motions tumbling down onto the ground. If the system can avoid evaluating such motions, the search for the optimal motion can be done much more efficiently.
3. It is not necessary to search manually the initial guess of the muscle-activation parameters that decide the motion, as the parameters that define the gait based on the AMPM are intuitive kinematic parameters such as the position and velocity of the COM. As explained in the Introduction, when using optimization methods based on forward dynamics, a good initial guess of the activation data must be provided in order to prevent the optimization process from getting caught into some local minima. This is quite time consuming, and the researcher needs special experience to determine the good set of muscle activations to generate realistic gait motion. In addition, as such process is done completely manually, the motion can become subjective to the researcher.

5. SUMMARY AND FUTURE WORK

In this study, we have proposed a new method to simulate the gait motion when muscles are deactivated. The method is based on the AMPM which is an enhanced model of the 3DLIPM that is often used in robotics to generate gait motion. Compared with previous methods, the proposed method is simple, and the computational cost is much smaller. As further work, we are planning to .t the AMPM model into human motion data, and simulate the effects of physiological disabilities and also the process of pathology for rehabilitation use.

6. REFERENCES

Anderson, F. C., Pandy, M. G., 2001. Static and dynamic optimization solutions for gait are practically equivalent. *Journal of Biomechanics* 34 (2), 153-161.

Crowninshield, R., Brand, R. A., 1981. A physiologically based criterion of muscle force prediction in locomotion. *Journal of Biomechanics* 14, 793-800.

Delp, S., 1990. Surgery simulation: A computer graphics system to analyze and design musculoskeletal reconstructions of the lower limb. Ph.D. thesis, Stanford University.

Hill, A, V., 1938. The heat of shortening and the dynamic constants of muscle. Proc. Roy. Soc., B 126, 136-195.

Kajita, S., Matsumoto, O., Saigo, M., 2001. Real-time 3d walking pattern generation for a biped robot with telescopic legs. Proceedings of the 2001 *IEEE International Conference on Robotics and Automatio* .

Kajita, S., Yamaura, T., Kobayashi, A., 1992. Dynamic walking control of a biped robot along a potential energy conserving orbit. *IEEE Transactions on Robotics and Automation* 8 (4). 8

Komura, T., Shinagawa, Y., 2001. Attaching physiological effects to motion-captured data. *Graphics Interface Proceedings* , 27-36.

Komura, T., Shinagawa, Y., Kunii, T. L., 2000. Creating and retargeting motion by the musculoskeletal human body model. *The Visual Computer* (5), 254-270.

Kudoh, S., Komura, T., 2003. C2 continuous gait-pattern generation for biped robots. Proceedings of *IEEE/RSJ International Conference on Intelligent Robots and Systems* , 1644-1650.

Neptune, R., 1999. Optimization algorithm performance in determining optimal controls in human movement analysis. *Journal of Biomedical Engineering* 124 (2), 249-252.

Neptune, R., Kautz, S., Zajac, F., 2001. Contributions of the individual ankle plantar flexors to support forward progression and swing initiation during walking. *Journal of Biomechanics* 34, 1387-1398.

Pandy, M., Zajac, F. E., 1991. Optimal muscular coordination strategies for jumping. *Journal of Biomechanics* 24, 1-10.

Pandy, M., Zajac, F. E., Sim, E., Levine, W. S., 1990. An optimal control model for maximum-height human jumping. *Journal of Biomechanics* 23 (12), 1185-1198.

Yamaguchi, G., Zajac, F. E., 1990. Restoring unassisted natural gait to paraplegics via functional neuromuscular stimulation: a computer simulation study. *IEEE Transactions on Biomedical Engineering* 37, 886-902.

Zajac, F. E., 1989. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *CRC Critical Reviews in Biomedical Engineering* 17 (4), 359-411.

Virtual environments as an aid to the design and evaluation of home and work settings for people with physical disabilities

O Palmon[1], R Oxman[2], M Shahar[3] and P L Weiss[4]

[1,3,4]Laboratory for Innovations in Rehabilitation Technology, Department of Occupational Therapy, University of Haifa, ISRAEL

[2]Faculty of Architecture and Town Planning, Technion, ITT, ISRAEL

[1]*palmon@research.haifa.ac.il,* [2]*arrro01@tx.technion.ac.il,* [3]*meir@shahar.name,* [4]*tamar@research.haifa.ac.il*

hw.haifa.ac.il/occupa/LIRT

ABSTRACT

One of the major challenges facing the professionals involved in the home modification process is to succeed in adapting the environments in a way that enables an optimal fit between the individual and the setting in which he or she operates. The challenge originates primarily from the fundamental characteristic of design - one can see and test the final result of home modifications only after they have been completed. The goal of this study was to address this problem by developing and evaluating an interactive living environments model, HabiTest, that will facilitate the planning, design and assessment of optimal home and work settings for people with physical disabilities. This paper describes the Habitest tool, an interactive model that has been implemented via an immersive virtual reality system which displays three-dimensional renderings of specific environments, and which responds to user-driven manipulations such as navigation within the environment and alteration of its design. Initial results of a usability evaluation of this interactive environment by users are described.

1. INTRODUCTION

People with disabilities are often severely limited in their ability to function independently in their homes or at work as a result of physical, cognitive or sensory environmental barriers (Iwarsson et al, 1998). Environmental modifications to an existing structure can eliminate these barriers thereby turning an environment from one that is disabling into one that is enabling. Such modifications aim to achieve an accessible environment that can be fully utilized by the user for its intended purposes (Nielsen and Ambrose, 1999) and are an inseparable part of the rehabilitation process in the western world (Gitlin, 1998; Iwarsson and Stahl, 2003). Yet, despite 30 years of experience and research, the gap between client needs and successful environmental modification remains large since adaptations to home or public settings present major challenges (Gitlin, 1998; Sundin and Medbo, 2003). In the home, the challenge relates to having sufficient knowledge and funds to make the necessary modifications. In public environments, the challenge is to succeed in making adaptations that will meet the needs of broad range of disabilities.

One of the central issues in environmental accessibility is the adaptation of environments to achieve an optimal fit between the individual and the setting in which he or she operates. Modifications to the physical environments in the home and at work to accommodate age and health related disability make these settings less challenging thereby facilitating the individual's ability to perform every day activities and occupations (Gitlin, 1998; Iwarsson et al, 1998).

One of the major challenges facing the professionals involved in the home modification process is to succeed in adapting the environments in a way that enables an optimal fit between the individual and the setting in which he or she operates. This challenge originates primarily from the fundamental characteristic of design - one can see and test the final result of home modifications only <u>after</u> they have been completed. Additional reasons for the limited success of environmental modifications include lack of suitable reliable and valid assessment tools and lack of common conceptual definitions between the professional parties involved in the environmental modification process (Gitlin, 1998; Iwarsson and Stahl, 2003; Steinfeld and Danford, 1999).

Virtual Reality (VR) has only recently begun to be applied to environmental modification. Eriksson et al, (2000) immersive, desktop simulation program was first tested through case studies (Eriksson and Johansson, 1996) and found to be a useful planning tool in encouraging communication and participation of all the people involved in the design process. This software was tested again more recently with a group of occupational therapy students and a group of people with physical disabilities (Eriksson, et al., 2000). Users view a typical living environment and are able to manipulate objects (e.g., furniture) within the environment. Although both groups enjoyed the opportunity to make modifications and to implement their own ideas for barrier-free design, the group with physical disabilities had difficulty in adopting a user-centred viewpoint.

More recent developments in software permit the creation of virtual environments that allow a designer to simulate three-dimensional (3-D) space and motion with a high degree of accuracy and detail (Moas, 2001; Kalisperis et al. 2002). Such programs enable the testing of architectural designs by the future tenants before the building has been constructed with the aim of minimizing the gap between a theoretical design and the actual end product (Moas, 2001). These tools enable the construction of and interaction within environments that are immersive, are ecologically valid, and can be graded with respect to their level of difficulty.

The overall goal of this study was to address this problem by developing and evaluating an interactive living environments model that will facilitate the planning, design and assessment of optimal home and work settings for people with physical disabilities. The first objective of this paper is to describe a tool that we have used to develop interactive environments that can be used to tests users' abilities to identify and modify accessibility barriers. This interactive model is implemented via an immersive virtual reality system which displays three-dimensional renderings of specific environments, and which responds to user-driven manipulations such as navigation within the environment and alteration of its design. A second objective is to present the initial results of a usability evaluation of this interactive environment by users.

2. THE DESIGN AND PROGRAMING OF HABITEST

We aimed to construct and evaluate a tool, known as HabiTest, that overcomes the inherent limitation of a posteriori design by providing a priori opportunities to verify the suitability of a proposed design for a particular user, using the option of virtual reality platform. This tool has been designed to address the needs of the environmental modification intervention process as well as the needs of a newly designed rendering.

2.1 Selection of simulation platform

The construction and simulation of these environments was carried out using EON Reality's (www.eonreality.com) tools, considered to be among the leading tools in the field of VR simulation. An environment used for our initial feasibility testing is shown in Fig. 1. In the figure are shown the three alternate points of view that are available to the user. These include a first person view, a third person view, and a bird's eye view. EON Reality's tools enable a rapid development of interactive 3-D environments that are easy to navigate in real time while performing accurate collision detection. Accurate collision detection (which was until recently available primarily in mechanical, non-interactive simulations) enhances our ability to gather relevant data from the simulation process. In previous generations of VR tools, the collision detection was limited to a bounding box. This bounding box was a rough approximation of the user's body contours and left out many of the fine details (such as curves, gaps and protrusions) which are needed to accurately represent the body.

EON Reality software not only enables identification of each collision, it also records their occurrences into a database. Moreover, auditory, visual and haptic feedback to the user prevents the attainment of positions that are physically invalid. That is, a user cannot navigate to a position where any part of his body nor an item associated with him (e.g., a wheelchair) is allowed to overlap with another object (e.g., wall, door, stair, table leg). The ecological validity of this simulation allows the user to identify corners or narrow passages that, although passable, would be difficult and inconvenient to navigate on a daily basis due to the number of moves and collisions they would necessitate.

2.2 *HabiTest Design Features*

The team that designed and programmed the Habitest consisted of two occupational therapists with expertise in environmental modification, assistive technology and virtual reality applications, two architects with expertise in spatial design and in three dimensional (3D) modelling and a programmer with expertise in virtual reality platform programming.

HabiTest must enable users to navigate independently within realistic virtual environments while allowing them to identify any barrier that blocks their ability to navigate or to perform tasks in these environments. HabiTest supports layered presentation. Since each object is associated with a particular layer, the display and elimination of any object is simple. Thus, an environment can be rapidly de-cluttered and 'what-if' testing is supported. For example, it is easy to query how navigation and obstacle identification are affected if any given object is added or removed. Removing layers of objects allows a user to perceive the effect this has on accessibility. It also allows users with cognitive or visual limitation to study the environment in an easier manner prior to testing it.

Figure 1. *Screen shots of a typical HabiTest environment. The larger central view is the main view and is "first person" (ego-centric). Two additional auxiliary views are available – "bird's eye" view (upper left corner) and "third person" side view (upper right corner). These views can be turned on and off as well as switched (i.e. the main view can be show bird's eye point of view while the top right auxiliary view can show the ego-centric view).*

A number of key features were considered to be essential for this tool. These include:

- real time rendering
- accurate collision detection of objects and walls
- low cost portable tool that could be readily implemented in clinical and home settings
- navigation that is easy and intuitive
- The avatar representing the user within the environment must have realistic characteristics (i.e., seated in a wheelchair with anthropometric measurement similar to an actual user.

In addition to the design requirements of ease of use and intuitiveness, the selection of a navigation tool was based on considerations of what would be the most appropriate hardware for use by people who drive an electric wheelchair since they are the initial targeted population for HabiTest. (Most electric wheelchairs are operated with joysticks.) We selected a force feedback joystick, specifically Logitech's Wingman and Microsoft's SideWinder. These both provide vibratory collision feedback to the user, enhancing the realism of the environment. Habitest can also be operated with other interface devices, most notably a standard mouse or even a head movement based mouse such as the TrakIR head mouse.

The size of the field of view (FOV) people need to navigate comfortably in an environment has certainly presented a challenge since healthy users who pilot tested HabiTest on a 17 inch monitor found the FOV presented to them to be too narrow for realistic and effective navigation. More recently we have improved the FOV in two different ways. The HabiTest environments may be projected onto a large surface using a video projector and/or several of the joystick buttons have been programmed to enable users to look around in a similar way people would turn their head to the right, left, up and down while exploring a new environment. The "looking around" option is available to users while standing in one place and while navigating in the environment.

The final design consideration is the provision of a means to perform tasks, and specifically to manipulate objects within the virtual environments. We have elected to provide a virtual hand which is controlled by the user. The first level of control is achieved by putting the hand at a fixed distance in front of the user's eyes; this means that when the user looks around, the hand will follow the look and will keep its position in the middle of the main view. Further control of the hand relates to its offset from the centre of the view. To simplify the interaction with the system, and reduce the different number of devices required (keyboard, joystick, mouse etc.), this offset is also controlled by the joystick. The user has to press and hold one of the joystick's buttons to activate this specific manipulation. After pressing the button, moving the joystick will affect this offset. To reset this offset, another button has to be pressed. The hand may be seen in the environment at any time.

3. INITIAL EVALUATION OF THE HABITEST BY USERS

3.1 Initial Usability Testing of HabiTest

The purpose of the initial usability testing was to verify the ease with which users are able to navigate within HabiTest. A group of eight female occupational therapists, all graduate students at the University of Haifa participated in the initial usability testing. All participants had worked clinically for a minimum of two years in different therapeutic settings dealing with developmental and learning disabilities of children. None of the participants had had any experience with environmental modifications, nor were they experienced in the playing of computer games that required navigation.

3.1.1 Procedure. The participants navigated within Habitest for 30 minutes while sitting in a computer lab, each at her own personal computer. The simulation environment was first presented to the participants by enslaving their monitors to the researcher's monitor. Each of the tool's features was demonstrated. A model of a one bedroom apartment with an open kitchen area, supplied by the EON interactive software (see Fig. 1) was modified to be used as a HabiTest environment for demonstration and testing. Following this short demonstration, each participant was given a trial period of ten minutes, to navigate within the environment freely using a mouse. During the trial period, participants were encouraged to request assistance when any operation was unclear or difficult. Upon completion of the trail period, the participants were asked to locate the bathroom, to drive into it and to exit it. They were given up to 20 minutes to perform this task from a starting point at the centre of the living room. This task was purposely selected as typical of a difficult navigation task since the location of the bathroom door together with its narrow doorway made the test a demanding one. Similarly, given the small size of the bathroom, the need to turn around within it in order to exit was also difficult. It has to be noted that the entrance of the bathroom was very difficult for a wheelchair but possible. Note that the collision detection feature was activated during the test, so that the participants were made aware of any "impossible" routes.

3.1.2 Results. Despite the fact that the participants had used the same environment during the practice trial, none of them were able to locate the bathroom within the first five minutes of the test. They appeared to be stymied by the narrow FOV and they expressed great frustration. The researcher then simplified the task by giving them instructions how to locate the bathroom but they were still required to enter and exit it. All participants were able to enter the bathroom but only five out of the eight were able to exit it within the 20 minute time limit. All participants experienced difficulty when trying to navigate within the narrow space of the bathroom. This was not surprising since the room was not accessible for a wheelchair user. The greatest concern expressed by the participants throughout the trial was their lack of awareness of their current location with respect to the apartment.

Similar difficulties were reported when a group of participants viewing the same HabiTest environment at a exhibition navigated within it using the joystick. As a result of the FOV difficulties observed during the two initial pilot tests we decided to expand the FOV via the two options described above, projection onto a large surface using a video projector and/or several of the joystick buttons have been programmed to enable users to look around in a similar way people would turn their head to the right, left, up and down while exploring a new environment.

3.2 Ongoing Usability Testing of the HabiTest

We are now testing the features of a new HabiTest environment navigated with a force feedback joystick. The participants include 15 volunteers, students at the University of Haifa, who are experienced users of a joystick for computer gaming; the extent of their familiarity with a joystick is recorded on a questionnaire that was designed for that purpose. The objective is to further test users' abilities to navigate within virtual environments when using a joystick.

3.2.1 Procedure. A new setting was created for the purpose of this testing. It is a rendered model representing the inner area of the main office section of the University's Dean of Students. This setting was selected since the physical premises have been made available to us for testing purposes. As shown in Fig. 2, this area contains a large central office and three additional small offices, all of which have doors leading to the central office. The area contains a lot of furniture making it difficult to navigate. Some locations are not accessible for a person who uses a wheelchair. Indeed, some of the furniture must be removed when students who use a wheelchair enter the premises. The rendered model of the Dean of Students central area is shown in Fig. 3. The figure on the left shows a "birds eye" view of the entire area and the figure on the right shows a “first person” view of one of the rooms showing its furniture.

Participants are shown how to use the joystick to navigate within HabiTest and given a trial period to practice the various operations. A 15 minute test period follows during which time participants explore the environment. Four noticeable (virtual) objects (e.g., red phone, plate with fruits, vase with red flowers) are placed in each of the four rooms. Once this period is over, the participants are given a map of the environment (printed on a paper) and asked to identify the location of each object. Participants are then requested to navigate to each object (within the virtual world) and to perform some activity with it (e.g., picking up an orange from the plate, picking up the phone). In order to be able to do so, they have to get within a reach range from the objects. During this test, participants are represented in the virtual environment as if they are sitting in a wheel chair. All actions are performed using one of the joystick’s programmed buttons thereby manipulating a “hand” that appears on the screen. All tests are timed and success or failure in performing all required activities is recorded. At the end of the testing periods the participants complete a usability questionnaire and a Presence questionnaire. The results of this testing will be reported.

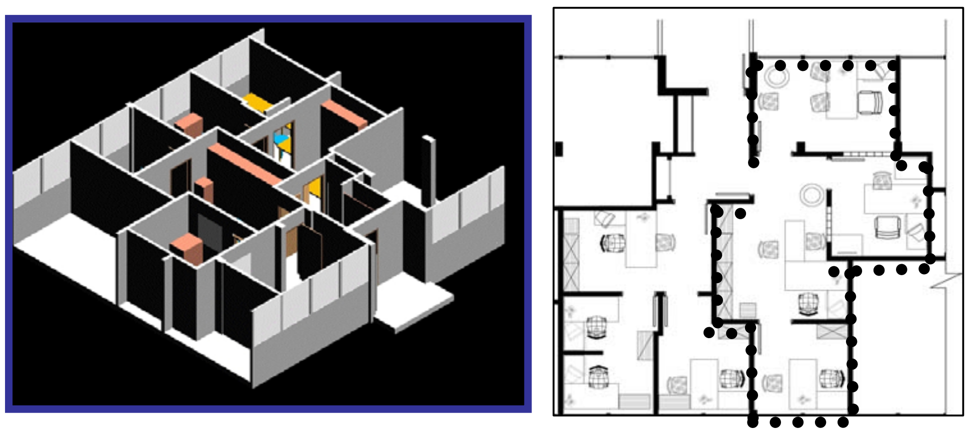


Figure 3. *Two views of the Dean of Students area as rendered by the EON Reality model. The figure on the left shows a “birds eye” view of the entire area. The figure on the right shows a blueprint of the same area (the heavy dotted line indicated the rendered area).*

3.3 *Future testing*

Our future plans include comparing performance within the HabiTest to that within an identical real environment in terms of: (1) ability to perform tasks and to navigate from one point to another (2) ability to identify accessibility barriers. Fifty male and female adults, aged 20 to 50 years, who use a powered wheelchair as a result of neurological (e.g., paraplegic spinal cord injury) or orthopaedic (e.g., amputation, arthritis) disability will participate in the study. They will all have had the disability for at least one year. The disability will be primarily of the trunk or lower extremity with sparing of the upper extremity. Prior to testing all participants will go through baseline measures that will test their functional level of personal independence and their ability to navigate a wheelchair in the real environment. Participants will use a joystick to navigate within the virtual environment. While navigating, participants will be confronted with different types of accessibility barriers (e.g. door is too narrow to pass through, table top is too low for a wheelchair to fit underneath). The participants will know that accessibility barriers exist in the environment; their task will be to identify what they are. A tester will observe the participants throughout the testing process. The EON interactive software will be used to render an interactive, 3D virtual environment, depicting real areas at the University of Haifa. The environments will have barriers that will make them inaccessible to people who use a wheelchair and will be programmed to enable the performance of specified tasks. Prior to testing, each participant will be instructed in the use of the HabiTest, software and hardware and given a trial period that will not be limited in time.

4. CONCLUSIONS

One of the major challenges facing the professionals involved in the home modification process is to succeed in adapting the environments in a way that enables an optimal fit between the individual and the setting in which he or she operates. We have begun to address this problem by developing and carrying out an initial evaluation of an interactive living environments model, HabiTest, that will facilitate the planning, design and assessment of optimal home and work settings for people with physical disabilities. This paper has described the design considerations taken into account while developing the Habitest tool. Several examples of interactive models have been pilot tested with users who are not disabled and their feedback has been used to improve the tool. We will soon commence testing HabiTest with a group of users who navigate with a wheelchair due to neurological or orthopaedic disability.

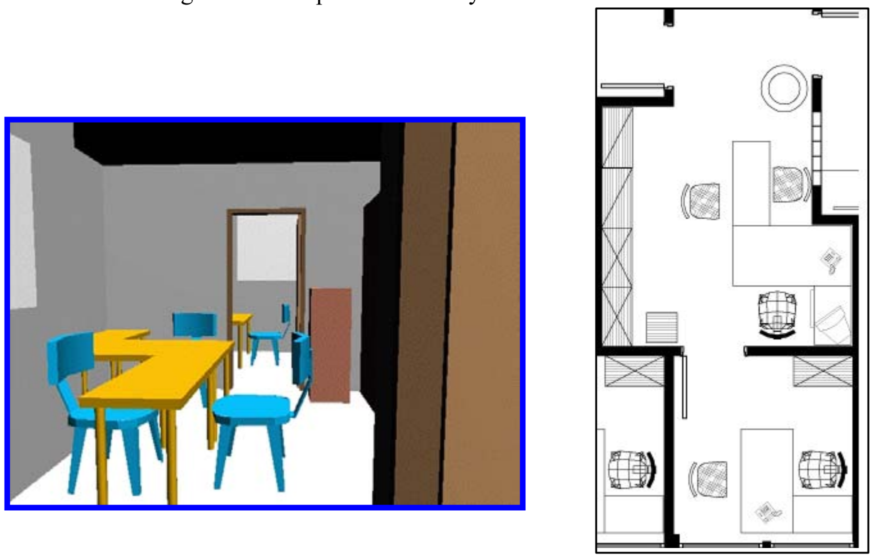


Figure 2. *Two images showing a part of the Dean of Students area used to test the HabiTest environment. The figure on the left displays a "first person" view of one of the rooms showing its furniture. The figure on the right shows the blueprint of the same area.*

Acknowledgements: We thank Prof. Skip Rizzo , Dr. Assaf Dvorkin and Dr Uri Feintuch for helpful feedback on HabiTest design and the user testing protocol. We also thank Noam Glikfeld from the Faculty of Architecture and Town Planning, the Technion, for producing the images shown in this paper and the 3D models that will be used in the testing of the HabiTest.

5. REFERENCES

A Moas (2001), *Virtual reality and architecture*, Retrieved from http://www.passing.com/pic/VRArchitecture.htp on June 26 2002. (Hebrew)

A Sundin and L Medbo (2003), Computer visualization and participatory ergonomics as methods in workplace design, *Hum Fac and Erg in Man,* **13**, 1, pp. 1-17.

C W Nielsen and I Ambrose (1999), Lifetime adaptable housing in Europe, *Tech & Dis,* **10**, pp. 11-19

J Eriksson, A Ek and G Johansson (2000), Design and evaluation of a software prototype for participatory planning of environmental adaptations, *Trans on Rehab Eng,* **8**, *1*, pp. 94-106.

J Eriksson and G Johansson (1996), dapting workplace and homes for disabled people using computer-aided design, *Industrial Ergonomics,* **17**, pp. 153-162.

L Kalisperis, G Otto, K Muramoto, J S Gundrum, R Masters and B Orland (2002), Connecting the Real and the Virtual, *Proc. of the virtual reality/space visualization in design education, ECAADE,* Warsaw, Poland, pp. 212-219.

L N Gitlin (1998), Testing home modification interventions: Issues of theory measurement, design and implementation, *Ann Rev of Gerontology and Geriatrics,* **18**, pp. 191-246

S Iwarsson (1999), The housing enabler: An objective tol for assessing accessibility. *BJOT*, **62**, *11*, pp. 491-497.

S Iwarsson, A Isacsson and J Lanke (1998), ADL Independence in the elderly population living in the community: The influence of functional limitations and physical environment demands, *OT Int,* **3**, pp. 52-61.

S Iwarsson and A Stahl (1999), Traffic engineering and occupational therapy: A collaborative approach for future directions. *SJOT*, **6**, pp. 21-28.

S Iwarsson and A Stahl (2003), Accessibility, usability and universal design – positioning and definition of concepts describing person-environment relationships, *Dis and Rehab*, **25**, *2*, pp. 57-66.

New accessibility model for Microsoft windows

R Haverty

Accessible Technology Group, Microsoft Corporation
1 Microsoft Way, Redmond, Washington 98052, USA

www.microsoft.com/enable

ABSTRACT

Microsoft® Windows® User Interface (UI) Automation is the new accessibility framework for Microsoft Windows and is intended to address the needs of assistive technology products and automated testing frameworks by providing programmatic access to information about the user interface. UI Automation will be fully supported in the Windows platform on "Longhorn" and will be the means of enabling automated testing and accessibility for all new forms of Windows user interface, including existing legacy controls.

1. INTRODUCTION

In 2003, Microsoft Corporation commissioned Forrester Research, Inc., to conduct a study to measure the potential market of people in the United States who are most likely to benefit from the use of accessible technology for computers http://www.microsoft.com/enable/research/. Accessible technology enables individuals to adjust their computers to meet their visual, hearing, dexterity, cognitive, and speech needs. It includes both accessibility options built into products as well as specialty hardware and software products (assistive technology products) that help individuals interact with a computer. Overall results show that 57% (74.2 million)[i] of computer users in the United States are likely or very likely to benefit from the use of accessible technology due to experiencing mild to severe difficulties or impairments http://www.microsoft.com/enable/research/computerusers.aspx.

With the current technology, assistive technology vendors (ATV) are required to use many different approaches to obtain and present information about the UI to the end user, thus spending an inordinate amount of time and resources on providing this basic information. With such a large percentage of users needing accessible information it is becoming increasingly important to make it easier for an ATV to programmatically obtain information about the UI.

The new accessibility model for Windows, UI Automation, is designed to provide a single reliable source of UI information to assistive technology products and automated test scripts. It provides programmatic access that allows automated tests to interact with the UI and allows assistive technology products to provide information about the user interface to their end users. UI Automation also provides means for manipulating the UI.

UI Automation has two main audiences: UI Automation providers and UI Automation clients. UI Automation providers are applications such as Microsoft Word, Excel, or third-party applications based on the Windows operating system. UI Automation clients are assistive technology applications, such as screen readers, screen enlargers, alternative input, or others. Automated test scripts can use UI Automation for automated testing and are also considered clients in the UI Automation framework.

This document includes information on the namespaces that the UI Automation framework uses, as well as information on the following UI Automation features: automation tree, UI Automation control patterns, UI Automation properties, UI Automation events, and UI Automation Input. Also included is information on security. This documentation for UI Automation is preliminary and is subject to change.

2. UI AUTOMATION NAMESPACES

The following table lists the namespaces used in the UI Automation framework, as well as the audience that uses each namespace.

Table 1. *UI Automation Namespaces and related audiences.*

Namespace	Audience used by
System.Windows.Automation	UI Automation clients (both assistive technology and automated test scripts) for finding automation elements, registering for events and working with control patterns.
System.Windows.Automation.Provider	UI Automation providers for implementing UI Automation on "Avalon" controls or applications.
System.Windows.Automation.InteropProvider	UI Automation providers for implementing UI Automation on non-"Avalon" frameworks such as Microsoft Win32®.
System.Windows.Automation.ComInteropProvider	UI Automation providers for implementing UI Automation on COM-based controls or applications.
System.Windows.Automation.Searcher	UI Automation clients for navigating the automation tree.

3. UI AUTOMATION TREE

Standard Windows programming has always exposed the relationship between elements in the user interface in a parent/child relational structure. UI Automation clients (assistive technologies and automated testing tools) view the UI elements on the desktop as a set of automation elements which are arranged in a tree structure. Automation elements implement a common class (AutomationElement) to enable consistent information, interaction, and a navigation model. UI Automation unifies disparate UI Frameworks such as Avalon, Trident, and Win32 so that code can be written against one API rather than several.

Within the automation tree there is a root automation element which represents the current desktop and whose children represent application windows on the desktop. Each of these childe elements can contain automation elements representing the UI elements comprising their UI, such as menus, buttons, toolbars, and list boxes. Each piece of UI can contain automation elements representing their content, such as menu items, or list items. Even a button, which does not contain any items, may have child automation elements which represent the basic UI components that comprise the button, such as text and rectangles.

It is important to note that the automation tree is not a fixed structure. For performance reasons it is built on demand starting with an automation element which the UI Automation client specifies.

1.1. Views of the Automation Tree

The automation tree can be filtered to create customized views of the tree which contains only those automation elements that are relevant for a particular client. This approach allows clients to customize the structure presented through UI Automation to their particular needs. A few default views are provided by the UI Automation framework, but clients can also define custom views if they need additional control.

1.2. Raw View

The raw view of the automation tree is the full tree of elements for which the desktop is the root. The raw view closely follows the native programmatic structure of an application and therefore is the most detailed view that is available. It is also the base on which the other views of the tree are built. Because this view depends on the underlying UI framework, the raw view of an Avalon button will have a different raw view than a Win32 button.

1.3. Control View

The Control View of the automation tree simplifies the assistive technology product's task of describing the UI to the end user and helping that end user interact with the application because it closely maps to the UI structure perceived by an end user.

The control view includes all elements from the raw view that an end user would understand as interactive or contributing to the logical structure of the control in the UI. Examples of elements that contribute to the logical structure of the UI but are not interactive themselves are item containers such as list view headers, toolbars, menus, and the status bar. Non-interactive elements used simply for layout or decorative purposes will not be seen in the control view. An example is a panel that was used only to layout the controls in a dialog but does not itself contain any information. Non-interactive elements that will be seen in the control view are graphics with information and static text in a dialog.

1.4. Content View

The content view of the automation tree contains elements that convey the true information in a user interface. For example, the values in a drop-down ComboBox will appear in the content view because they represent the information being used by an end user. In the content view, a ComboBox and ListBox are both represented as a collection of items where one, or perhaps more than one, item can be selected. The fact that one is always open and one can expand and collapse is irrelevant in the content view because it is designed to show the data, or content, that is being presented to the user.

1.5. Custom Views

The UI Automation framework also allows a UI Automation client to create a custom view of the automation tree by specifying the desired match conditions and scoping information. This also allows UI Automation clients to build their own interaction models for the application using just the data that they need.

1.6. Automation Tree Structure Example

The following example compares the control view and content view of the automation tree for the same application: Wordpad.

Table 2. *Comparison of two Automation Tree structures*

Automation Tree (Control View)	**Automation Tree (Content View)**

The *Control* view of WordPad shown from the Desktop has the following automation tree structure:
• Desktop
  o Window “Notepad”
    ▪ TitleBar “Notepad”
      • SystemBar
        o MenuItem
      • Button AutomationId = “Minimize”
      • Button AutomationId = “Maximize”
      • Button AutomationId = “Close”
    ▪ MenuBar ““
      • MenuItem “File”
      • MenuItem “Edit”
    ▪ ToolBar ““
      • Button “New”
      • Button “Open”
    ▪ Text ““
    ▪ StatusBar
      • Edit
      • Edit
	The *Content* view of WordPad shown from the Desktop has the following automation tree structure:
	• Desktop
	  o Window “Notepad”
	    ▪ MenuBar ““
	      • MenuItem “File”
	      • MenuItem “Edit”
	    ▪ ToolBar ““
	      • Button “New”
	      • Button “Open”
	    ▪ Text ““
	    ▪ StatusBar
	      • Edit
	      • Edit

4. UI AUTOMATION CONTROL PATTERNS

UI Automation uses control patterns to express the functionality contained in each control. UI Automation differentiates between what a user would call the control and what can be programmatically done with the control by using control patterns to express only functionality, separate from the type or name of that control.

UI Automation providers implement control pattern interfaces on UI elements. For Avalon controls, control pattern interfaces are found in the System.Windows.Automation.Provider namespace and have names that include the suffix “Provider” (for example, IScrollProvider and IInvokeProvider). For non-Avalon controls, control pattern interfaces are found in the System.Windows.Automation.InteropProvider namespace and have names that include the suffix “InteropProvider” (for example, IScrollInteropProvider and IInvokeInteropProvider).

Clients access methods and properties of control pattern classes and use them to access information about a UI element, or to manipulate the UI. These control patterns classes are found in the

Systems.Windows.Automation namespace and have names that include the suffix "Pattern" (for example, InvokePattern and SelectionPattern).

1.7. Control Pattern Components

Control patterns define the structure, methods, properties, and events supported by a control:

- The structure includes the parent, child, and sibling relationships of elements for that control pattern.
- The methods provide the ability to programmatically manipulate the control.
- The properties and events provide rich information and notifications relevant for that control.

Control patterns relate to UI as interfaces relate to COM objects. In COM, you can query an object to ask what interfaces it supports, and then use those interfaces to access functionality. In UI Automation, clients can ask a control which patterns it supports and then interact with the control through the properties, methods, events, and structure of the supported control patterns. For example, providers implement IScrollProvider for a multi-line edit box. When a client detects that a UI element supports ScrollPattern, it can use the properties, methods, and events from that class to gather scroll-specific information or programmatically scroll its content to a new position.

1.8. Standard UI Controls and Their Control Patterns

Controls can support zero or more control patterns. For example:

- The image control does not support any control patterns.
- The button control supports InvokePattern to correspond to the functionality that it can be clicked.

To define the full set of functionality for a control, providers implement multiple control patterns. The following table shows more examples of standard controls and the control patterns they support.

Table 3. *Controls and their Control Patterns.*

Control Type	Relevant Control Patterns
Button	Invoke
CheckBox	Toggle
ComboBox	ExpandCollapse or Selection
Edit	Value, Text
ListBox	Selection
ListItem	SelectionItem
Tree	Selection
TreeItem	SelectionItem, ExpandCollapse

1.9 Control Patterns

Table 4 lists some of the key control patterns and their classes and interfaces. Note that interfaces are designated for "Avalon" providers (for example, IInvokeProvider) and non-"Avalon" providers (for example, IInvokeInteropProvider).

5. UI AUTOMATION PROPERTIES

UI Automation properties are a set of standard properties that expose information that is important to assistive technologies. Frequently, this information is exposed differently for each UI framework. Table 5 shows how one standard UI Automation property maps to multiple property names in other UI frameworks. By implementing UI Automation, providers map unique UI framework properties to standard UI Automation properties. When this done, it allows UI Automation clients to query for property information using one API call for a UI Automation property.

Table 4. *Control Patterns and their Classes and Interfaces*

Control Pattern	Client-Side Class	Provider-Side Interfaces
ExpandCollapse	ExpandCollapsePattern	IExpandCollapseProvider; IExpandCollapseInteropProvider
Grid	GridPattern	IGridProvider; IGridInteropProvider
GridItem	GridItemPattern	IGridItemProvider; IGridItemInteropProvider
Invoke	InvokePattern	IInvokeProvider; IInvokeInteropProvider
MultipleView	MultipleViewPattern	IMultipleViewProvider; IMultipleViewInteropProvider
RangeValue	RangeValuePattern	IRangeValueProvider; IRangeValueInteropProvider
Scroll	ScrollPattern	IScrollProvider; IScrollInteropProvider
Selection	SelectionPattern	ISelectionProvider; ISelectionInteropProvider
SelectionItem	SelectionItemPattern	ISelectionItemProvider; ISelectionItemInteropProvider
Sort	SortPattern	ISortProvider; ISortInteropProvider
Table	TablePattern	ITableProvider; ITableInteropProvider
TableItem	TableItemPattern	ITableItemProvider; ITableItemInteropProvider
Text	TextPattern	ITextProvider; ITextInteropProvider
Toggle	TogglePattern	IToggleProvider; IToggleInteropProvider
Transform	TransformPattern	ITransformProvider; ITransformInteropProvider
Value	ValuePattern	IValueProvider; IValueInteropProvider
Window	WindowPattern	IWindowProvider; IWindowInteropProvider
Zoom	ZoomPattern	IZoomProvider; IZoomInteropProvider

Table 5. *Mapping UI Automation Properties to Other UI Frameworks.*

UI Automation Control Type	UI Framework	Framework Property	UI Automation Property
Button	Avalon	Content	NameProperty
Button	Win32	Caption	NameProperty
Image	Trident/HTML	ALT	NameProperty

6. UI AUTOMATION EVENTS

UI Automation offers an event mechanism similar to WinEvents in the current Windows platform. However, unlike WinEvents, UI Automation's events are not based on a broadcast mechanism. UI Automation clients register for specific event notifications and can request that specific UI Automation properties and control pattern information be passed into their event handler. This provides a much more powerful mechanism than WinEvents because clients make fewer calls to retrieve the information they require, which results in a fewer cross-process calls, and therefore better performance. UI Automation provides event notifications for logical structure changes, control pattern changes, focus changes, property changes, and multimedia events.

7. UI AUTOMATION INPUT

In addition to using control patterns to manipulate the UI, UI Automation provides an Input class as a way to automate keyboard or mouse input. This class is a simple wrapper around SendInput that may, in a future release, be replaced by a more powerful input model (designed in conjunction with IME, Tablet, and Cicero teams).

8. SECURITY AND UI AUTOMATION

The security model for UI Automation is based on only granting access that is needed at the current time. Programmatically accessing and manipulating the user interface (UI) - functionality available in many UI Automation programming elements - requires specific security permissions. These UI Automation programming elements do not work inside a default secure execution environment. This applies to the following groups of methods:

- Methods that access information about a UI element, such as property values.
- Methods that write or modify the UI by using Control Patterns, such as AddToSelection.
- Methods that provide keyboard or mouse input. These methods are in the Input class.

UI Automation clients need to have the appropriate AutomationPermission when using these programming elements. The following table summarizes the AutomationPermission that UI Automation clients can use and provides examples of methods that would require the permission:

Table 5. *AutomationPermssion for UI Automation Clients*

Action of Method	AutomationPermissionFlag	Example
Access information about the UI	Read	GetPropertyValue
Write to or modify the UI	Write	Control pattern methods that manipulate the UI, such as AddToSelection or Invoke.
Send mouse or keyboard input	Input	Methods from the Input class, such as MoveToAndClick or SendKeyboardInput.

9. CONCLUSIONS

UI Automation is a key part of the new accessibility model for Windows, gathering information about and interacting with the UI. Adoption of this technology will improve product quality for Windows applications and reduce the time to market for assistive technology products. Additionally, by implementing UI Automation, ATVs reduce the resources invested in obtaining UI information allowing them to improve and expand on the products that they offer.

10. REFERENCES

UI Automation specific section of the Longhorn Software Development Kit (SDK):

http://longhorn.msdn.microsoft.com/?//longhorn.msdn.microsoft.com/lhsdk/accessibility/overviews/uiaccess_ovw.aspx,

Longhorn SDK Home page: http://longhorn.msdn.microsoft.com/

Longhorn Developer Center home page: http://msdn.microsoft.com/longhorn/

Microsoft research results, “The Wide Range of Abilities and Its Impact On Computer Technology”:

http://www.microsoft.com/enable/research/

Microsoft Accessibility home page: http://www.microsoft.com/enable/

[i] Study Commissioned by Microsoft Conducted by Forrester Research, Inc. 2004.

Blind Persons' Acquisition of Spatial Cognitive Mapping and Orientation Skills Supported by Virtual Environment

O Lahav[1] and D Mioduser[2]

School of Education, Tel Aviv University,
Ramat Aviv, Tel Aviv, ISRAEL

[1]lahavo@post.tau.ac.il, [2]miodu@post.tau.ac.il

[1]muse.tau.ac.il/orly/, [2]muse.tau.ac.il/david/

ABSTRACT

Mental mapping of spaces, and of the possible paths for navigating these spaces, is essential for the development of efficient orientation and mobility skills. Most of the information required for this mental mapping is gathered through the visual channel. Blind people lack this crucial information and in consequence face great difficulties (a) in generating efficient mental maps of spaces, and therefore (b) in navigating efficiently within these spaces. The work reported in this paper follows the assumption that the supply of appropriate spatial information through compensatory sensorial channels, as an alternative to the (impaired) visual channel, may contribute to the mental mapping of spaces and consequently, to blind people's spatial performance. The main tool in the study was a virtual environment enabling blind people to learn about real life spaces, which they are required to navigate.

1. INTRODUCTION

The ability to explore unknown spaces independently, safely and efficiently is a combined product of motor, sensory and cognitive skills. Normal exercise of this ability directly affects individuals' quality of life. Mental mapping of spaces, and of the possible paths for navigating these spaces, is essential for the development of efficient Orientation and Mobility (O&M) skills. Most of the information required for this mental mapping is gathered through the visual channel (Lynch, 1960). People who are blind lack this information, and in consequence they are required to use compensatory sensorial channels and alternative exploration methods (Jacobson, 1993). The research reported here is based on the assumption that the supply of appropriate spatial information through compensatory sensorial channels, as an alternative to the (impaired) visual channel, may help to enhance blind people's ability to explore unknown environments (Mioduser, in press).

Research on O&M skills of people who are blind in known and unknown spaces (e.g., Passini & Proulx, 1988; Ungar, Blades & Spencer, 1996) indicates that support for the acquisition of spatial mapping and orientation skills should be supplied at two main levels: perceptual and conceptual. At the perceptual level, the deficiency in the visual channel should be compensated by information perceived via other senses. The haptic, audio and smell channels become powerful information suppliers about unknown environments. For blind individuals, haptic information is commonly supplied by the white cane for low-resolution scanning of the immediate surroundings, by palms and fingers for fine recognition of object form, texture and location, and by the feet regarding navigational surface information. The auditory channel supplies complementary information about events, the presence of other people (or machines or animals) in the environment, or estimates of distances within a space (Hill, et al., 1993).

As for the conceptual level, the focus is on supporting the development of appropriate strategies for an efficient mapping of the space and the generation of navigation paths. Research indicates that people use two main scanning strategies: route and map strategies. Route strategies are based on linear (and therefore sequential) recognition of spatial features, while map strategies, considered to be more efficient than the former, are holistic in nature, comprising multiple perspectives of the target space (Fletcher, 1980; Kitchin & Jacobson, 1997). Research shows that people who are blind use mainly route strategies when recognizing and navigating new spaces (Fletcher, 1980).

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

Advanced computer technology offers new possibilities for supporting rehabilitation and learning environments for people with disabilities (e.g., sensorial, physical, mental, and learning disabilities) (Schultheis & Rizzo, 2001). It has also been used for rehabilitation for blind people; in particular, Virtual Environment (VE), which includes haptic interface technology, enables blind individuals to expand their knowledge as a result of using artificially made reality through haptic and audio feedback. Research on the implementation of haptic technologies within virtual navigation environments has yielded reports on its potential for supporting rehabilitation training with sighted people (Giess, Evers, & Meinzer, 1998; Gorman, et al., 1998), as well as with people who are blind (Jansson, et al., 1998; Colwell, Petrie & Kornbrot, 1998). Related research on the use of haptic devices by people who are blind, includes the following: identification of texture and object shape (Sjotrom & Rassmus-Grohn, 1999), mathematical learning environment and exploring of mathematical graphs (Karshmer & Bledsoe, 2002), and construction of cognitive maps (Lahav and Mioduser, 2000; Semwal & Evans-Kamp, 2000). In our previous research, we have shown that the use of VE technology helped people who are blind in exploring an unknown novel room (Lahav & Mioduser, 2003; 2004).

The research reported in this paper follows the assumption that the supply (via the technology) of compensatory perceptual and conceptual information may contribute to blind persons' cognitive mapping of spaces. To examine the above assumption we developed a multisensory-virtual-learning-environment (MVLE) and studied the exploration process of an unknown space by blind subjects using this VE. Their performance was compared to that of a control group of blind people who explored directly the real environment simulated in the MVLE. The main research questions of this study were:

1. What characterizes a blind person's exploration process of an unknown environment using a VE?
2. Does walking in the VE contribute to the construction of a cognitive map of the unknown space?
3. How does this cognitive map contribute to the blind person's O&M performance in the real environment?

2. THE HAPTIC VIRTUAL ENVIRONMENT

For the study we developed a VE simulating real-life spaces. This VE comprises two modes of operation:

2.1 Developer/Teacher Mode

The core component of the developer mode is the VE editor (Figure 1). This module includes three tools:

(a) *3D environment builder* – using this builder the developer defines such physical characteristics of the space as size and form of the room, type and the size of objects (e.g., doors, windows, furniture pieces) and their location. Although the environment builder is based on a 3D editor, for this research we used a 2D environment only.

(b) *Haptic feedback output editor* – this editor permits the developer to attach force-feedback effect to all components in the VE. The main haptic feedback is the kinesthetic force that the user feels through the Force Feedback Joystick (FFJ). The user feels the variation in texture and friction of the virtual component simulated in the VE.

(c) *Audio feedback output editor* - the audio feedback was used to provide the user with a comfortable channel carrying descriptive information. Our VE includes three types of audio feedback: (i) the intentional tapping or accidental bump with the force feedback joystick on one of environments' components (e.g., doors, walls, or box) initiates the audio feedback indicating the object's identity; (ii) the computer automatically generates an audible alert when approaching obstacles' corner; (iii) during navigation, the provision of footstep sounds and echoes increases not only the reality to the blind user, but also the sense of actual scale. The sound interval of the footsteps shows the speed of the navigation; and the user's stride-length is the benchmark for distance in the virtual scene.

2.2 Learning Mode

The learning mode, within which the user works, includes two s interfaces:

(a) *User interface* – the user interface consists of the VE that simulates real rooms and objects to be navigated by the users using the FFJ (Figure 2).

(b) *Teacher interface* - the teacher interface includes several features that serve teachers during and after the learning session. On-screen monitors present updated information on the user's navigation performance, such as position, or objects already reached. An additional feature allows the teacher to record the subject's navigation path, and replay it to analyze and evaluate the user's performance (Figure 3).

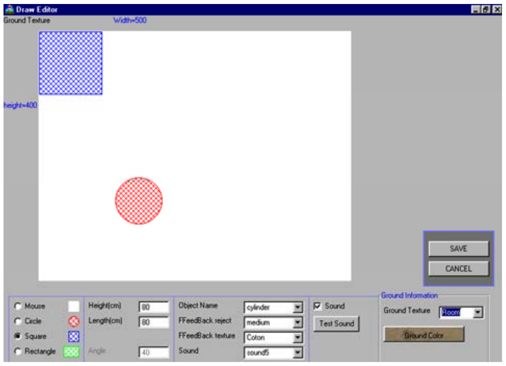


Figure 1. *3D environment editor.*

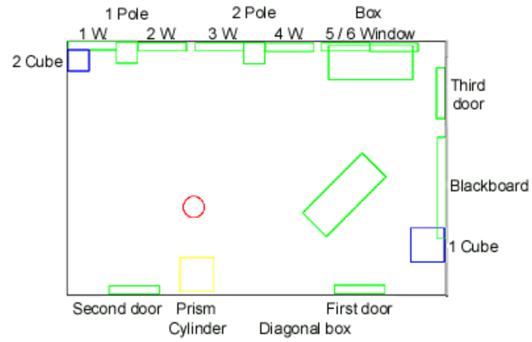


Figure 2. *The VE representation of the target space.*

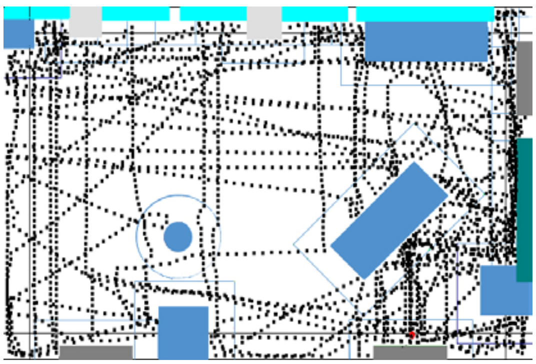


Figure 3. *Subject's navigation path.*

Figure 4. *The real space.*

3. METHOD

3.1 Subjects

The study included 31 subjects who were selected on the basis of the following seven criteria: (i) total blindness; (ii) at least 12 years old; (iii) not multi-handicapped; (iv) received O&M training; (v) Hebrew speakers; (vi) onset of blindness at least two years prior to the experimental period and (vii) comfortable with the use of computers. The subjects' age range was 12-70 years, mostly adults in the age range of 24-40. We defined two groups that were similar in gender, age and age of vision loss (congenitally blind or late blind): The experimental group, including 21 subjects who explored the unknown space by means of the VE, and the control group, including ten subjects who explored directly the real unknown space.

3.2 Research Instruments

Seven main instruments served the study; the last five instruments were developed for the collection of quantitative and qualitative data. The research instruments were:

(a) *The Unknown Target Space* - The space to be explored, both as real physical space (see Figure 4) and as a virtual space in the VE (see Figure 2), was a 54-square-meter room with three doors, six windows and two columns. There were seven objects in the room, five of them attached to the walls and two placed in the inner space.

(b) *Exploration Task* - Each participant was asked individually to explore the room, without time limitations. The experimenters informed the subjects that they would be asked to describe the room and its components at the end of their exploration.

(c) *Orientation and Mobility Questionnaire* - The questionnaire comprised 46 questions concerning the subjects O&M ability indoors and outdoors, in known and unknown environments. Most of the questions were taken from O&M rehabilitation evaluation instruments (e.g., Dodson-Burk & Hill, 1989; Sonn, Tornquist & Svensson, 1999).

(d) *Interview* - An open interview used for the subjects' verbal description of his exploration in the unknown environment.

(e) *Observations* - For recording the participant's exploration, we used video-recorder and notebook. Their navigation process and audio remarks in the VE and the real space were recorded during the tasks. The information from these recordings was combined with the computer log recording.

(f) *Computer Log* - The Log allowed the researchers to track the user's learning and exploration process in the MVLE, as regards to their exploration strategies, distances traversed, duration, switch of strategies and breaks.

(g) *Evaluation and Coding Schemes* - These instruments served the experts' analysis of the participant's O&M skills and capabilities and his or her acquaintance process with the new space.

3.3. Procedure

All subjects worked and were observed individually. The study was carried out in five stages: (i) evaluation of the subjects' initial O&M skills, using the O&M questionnaire; (ii) familiarization with the VE features and how to operate the FFJ (the experimental group); (iii) subjects' exploration of the unknown space: the experimental group explored the space using the VE, while the control group explored the real environment directly; (iv) following the exploration task the subjects were asked to give a verbal description of the environment, and to construct a scale model of it; (v) performing O&M tasks in the real target space: Target-object task (the user will be ask to find an object in the space); Perspective-taking task (the user will enter the room by a different entrance and asked to find an object). In the last four stages all subjects' performances were video-recorded.

4. RESULTS

Research Question 1: *What characterizes a blind person's exploration process of an unknown environment using a VE?*

Significant differences were found between the experimental group and the control group concerning the characteristics of the exploration process. These differences are related to four variables: the total duration of the exploration, the total distance traversed, the sequence of main strategies implemented and the number of pauses made while exploring the unknown space. Data in Table 1 show significant differences between the experimental and the control groups in that the experimental group took more breaks during their exploration tasks.

Table 1. *Short and long breaks.*

Group	Long breaks	Short breaks
Experimental group (n=21)	17	81
Control group (n=10)	6	13
	*	**

* p<.05; ** p<.001

The subjects in both groups implemented similar exploration strategies, mostly based on the ones they use in their daily navigation in real spaces (for example: "perimeter", i.e. walking along the room's walls and exploring objects attached to the walls; "grid", i.e.. exploring the room's inner-space). However, an interesting additional finding is that several subjects in the experimental group developed a few new strategies while working within the VE. Those strategies could be generated only within the VE, representing an important added value of the work with the computer system. Although no substantial difference between groups was observed as regards the types of strategies used, significant difference was found concerning the frequency of use of the strategies, and distance traversed using each strategy. Data in Table 2 indicate that the strategy most frequently used by the experimental group was grid, followed by the perimeter strategy. In contrast, the control group preferred to explore the room's perimeter, and next, to use the object-to-object strategy. A detailed presentation of the findings of the exploration stage can be found in Lahav & Mioduser, (2004).

Table 2. *Exploration strategies, frequency and length.*

Exploration patterns	Experimental group (n=21)		Control group (n=10)	
	Frequency	Length of the path (In meters)	Frequency	Length of the path (In meters)
Perimeter	86	53.9	28	14.6
Grid	116	26.3	9	.97
Object-to-object	22	7.8	14	2.3
Points of reference	50	26.6	13	1.8
New strategies	18	18.2	--	--
Sum	292	132.8	64	19.67
Mean	14	6.3	6.4	1.9

Research Question 2: *Does walking in the VE contribute to the construction of a cognitive map of the unknown space?*

After completing the exploration task the subjects were asked to give a verbal description and to construct a model of the environment. Four variables of the subjects' verbal and physical representations were examined: room size, room shape, structural features and components' location. The control group subjects (who explored the real space directly) performed better in verbally describing the rooms' size ($\chi2(2)=9.07$; $p<0.05$) and the rooms' shape ($\chi2(2)=7.02$; $p<0.05$). The subjects from the experimental group performed better in describing the structural components ($t(28)=4.63$; $p<0.001$) and their location ($t(29)=2.85$; $p<0.001$).

Most subjects in both groups constructed an appropriate model of the room and its components. Data related to aspects of the subjects' reference to structural components and objects in the environment are shown in Table 3. Significant differences between the two research groups were observed in six variables. The data demonstrated that the information resolution of the components of the cognitive map built by subjects of the experimental group was finer in detail than the map built by subjects of the control group.

Table 3. *Cognitive map construction - verbal description and model construction.*

	Verbal description			Model construction		
	Experimental group (n=21)	Control group (n=10)		Experimental group (n=21)	Control group (n=10)	
Room size	--	50%	*			
Room shape	15%	60%	*			
Structures' components	46	16	**			
Estimation Structures' components location	20	7	*			
Model structure				95%	100%	
Objects	79	53	*	83	48	*
Estimation object location	60	40		50	28	
Placing object in the room	2	3		60	29	
Estimating object size	0.7	10		41	27	

*p<.05; **p<.001

The experimental group's representation was more specific and elaborate, in both verbal description and model construction. For example, 29% of experimental group subjects placed all seven objects located in the environment in their model, and 43% placed six objects. In contrast, none of the control group subjects placed all seven objects in their model and only 30% placed six objects.

The findings for the second question indicate that the experimental group subjects constructed fairly complex cognitive maps of the unknown space, as reflected in their verbal and physical descriptions. These maps comprise multiple layers, including the structural layer (referring to the overall configuration and dimensions of the room), the compositional layer (in relation to the identification of inner components and their arrangement in space), and the relational layer (location of objects relative to each other, or distances among objects). A procedural component complements the previous layers in the form of strategies for exploration/recall of the target space (e.g., perimeter, object-to-object). The learning process within the VE, by its unique features, supported the construction of a knowledge-rich model at all its different layers. A detailed presentation of the findings of the cognitive map construction stage can be found in Lahav & Mioduser (in press).

Research Question 3: *How does this cognitive map contribute to the blind person's O&M performance in the real environment?*

After the construction of the cognitive map, the subjects were asked to perform two orientation tasks in the real space. It should be recalled that the experimental group subjects entered the real space for the first time to perform the tasks, and were not given the option to first explore the room (initial exploration was accomplished in the VE only). Five variables were examined: successful completion of the tasks; use of direct paths to the target location; time spent on task; number and duration of breaks (short breaks and long breaks) and total length of the path (see Table 4). Most of the subjects of the experimental group successfully performed both orientation tasks in the real space. Significant difference was found between the groups in the subjects' performance in the target-object task. Most subjects of the experimental group successfully performed the target-object task while choosing a more direct and shorter path than the control group subjects; about half of the experimental group subjects choose a straight walking path and the "Object-to-Object" strategy. When examining the perspective-taking task, most subjects of the experimental group successfully performed the task in shorter time and path length, 50% using the "perimeter" strategy.

Table 4. *Performance in the real environment.*

	Target-object task			Perspective-taking task	
	Experimental group (n=21)	Control group (n=10)		Experimental group (n=21)	Control group (n=10)
Success (%)	81%	40%	*	71%	60%
Direct path (%)	67%	20%	**	34%	30%
Time (Seconds)	66	118		153	191
Short breaks (mean)	3	6		3	5
Long breaks (mean)	1.5	2.7		1.5	3
Length of the path	28	47	***	86	95

* $\chi2(2)=7.02$; $p<0.05$; ** $\chi2(3)=8.20$; $p<0.05$; *** $p<0.05$

The results are clearly indicative of the contribution of VE learning to the participants' anticipatory mapping of the target space and consequently to their successful performance in the real space. Moreover, they show that such a mapping resulted in greater capability of the subjects of the experimental group in performing the real-space tasks.

5. CONCLUSIONS

5.1 Exploration of an unknown environment using a VE

Walking in the VE gave participants a stimulating, comprehensive and thorough acquaintance with the target space. The high degree of compatibility between the components of the VE and of the real space on one hand and the exploring methods supported by the VE on the other, contributed to the users' relaxed and safe walking. These features also enabled participants to implement exploration patterns they commonly used in real spaces, but in a qualitatively different manner. The use of "real exploration strategies" in VEs was reported in previous studies on spatial performance by sighted participants (Darken & Peterson, 2002; Witmer, et al., 1996). But this study's VE participants applied the known strategies in novel ways; for example, they preferred to explore the inner part of the room first and only then its boundaries (in contrast with the exploration patterns described by Jacobson, 1993). Moreover, the VE participants created new exploration strategies, such as the one simulating walking with a long cane enabling them to walk the perimeter of the room and at the same time to explore its corresponding inner areas – a strategy only possible within the VE. Operation features of the VE (e.g., the game-like physical interface, multiple types of feedback) contributed to participants' confident performance with the system while exploring the unknown space. As a result, the exploration process showed interesting qualities concerning spatial, temporal, and thinking-related aspects. Examples of spatial and temporal qualities are the range of scanning strategies implemented, the inclusion of a large number of long and short breaks, or the time spent in examining the space. Concerning thinking-related aspects of the process, interesting examples were the long breaks made by the participants with the aim to reflect on the exploration steps or to memorize data concerning an explored area, or the use of "virtual drawing" of spatial features under examination on the table's surface as reinforcement aid.

5.2 Construction of a cognitive map of the unknown space as a result of learning with the VE

Participants in the experimental group were able to construct complex maps of the unknown space while working with the VE, prior to their acquaintance with the real space. As a result of their intensive interaction with the components of the virtual learning environment, the users were exposed to a wide range of haptic and audio feedbacks. This information allowed them to devote most of their attention and resources to the consolidation of the structural, compositional and relational aspects of the space's overall map. In addition, it seems that the participants developed particular perspectives of the space, and strategies for approaching it, as a result of the features of the VE (e.g., the tendency to describe the space from the perimeter to the inner space, in a whole and holistic manner). In contrast, we found that the exploration of the real space contributed to the control group's ability to estimate the objects' size and distances among them, functions not supported yet in the VE.

5.3 Performance of O&M tasks as a result of learning with the VE

The first real space walking experience of most subjects in the experimental group was a confident and resolved one. It was noticeable that this walking was based solely on spatial knowledge acquired as a result of their acquaintance with the room in the VE. We found many evidences of the robustness of the constructed map and its contribution to the subjects' performance. One example is their frequent use of the "Object-to-Object" strategy while accomplishing the tasks. Previous research (e.g., Golledge et al., 1996; Hill et al., 1993) reported that successful navigators among people who are blind make recurrent use of this strategy. The frequent use of this strategy by participants in the experimental group is indicative of the holistic nature of their inner representation of the space, allowing them to construct efficient navigation paths based on isolating subsets of objects and their relative location. This internal representation represents a powerful tool for guiding the secure navigation in the real space immediately after entering it, and for locating all spatial components required to perform the task in the shortest possible time and ambulation path.

5.4 Current constraints and future implications

In this first attempt to examine an exploration task, the cognitive mapping process, and its effect on actual performance we had to stay within the limits of the first stage of a more comprehensive research agenda. For example, in selecting the characteristics of the target space in this first stage we focused on a closed space without complicated topographical traits (which represent a complete set of additional variables that might lead to different results). Based on this first stage our research agenda includes the modeling of increasingly open and complex spaces (e.g. university campus, museum, public building), the offering of additional exploration tools, and the study of overall exploration and mapping strategies by people who are blind who recurrently use the VE for different spaces. We believe these studies are of theoretical and practical value for (a) training and rehabilitation processes requiring the acquisition of orientation and mobility skills and strategies, and (b) learning processes of subjects involving spatial information, by congenital and late blind people.

Acknowledgements: The study presented here is partially supported by grants from the Israeli Ministry of Education, Microsoft Research Ltd., and Israel Foundation Trustees.

6. REFERENCES

Colwell, C., Petrie, H., & Kornbrot, D. (1998). *Haptic virtual reality for blind computer users*. Paper presented at the Assets '98 Conference, Los Angeles, CA. Available in: http://phoenix.herts.ac.uk/sdru/pubs/VE/colwell.html

Darken , R.P., & Peterson, B. (2002). Spatial orientation, wayfinding and representation. In K. M. Stanney (Ed.), *Handbook of virtual environments design, implementation, and applications* (pp. 493-518). Hillsdale, NJ: Erlbaum.

Dodson-Burk, B., & Hill, E.W. (1989). Preschool orientation and mobility screening. *A publication of division IX of the association for education and rehabilitation of the blind and visually impaired*. New York, NY: American Foundation for the Blind.

Fletcher, J.F. (1980). Spatial representation in blind children 1: development compared to sighted children. *Journal of Visual Impairment and Blindness, 74* (10), 318-385.

Giess, C., Evers, H., & Meinzer, H.P. (1998). *Haptic volume rendering in different scenarios of surgical planning*. Paper presented at the Third Phantom Users Group Workshop, M.I.T. Cambridge, MA.

Golledge, R., Klatzky , R., & Loomis, J. (1996). Cognitive mapping and wayfinding by adults without vision. In J. Portugali (Ed.), *The construction of cognitive maps* (pp. 215-246). The Netherlands: Kluwer.

Gorman, P.J., Lieser, J.D., Murray, W.B., Haluck, R.S., & Krummel, T.M. (1998). *Assessment and validation of force feedback virtual reality based surgical simulator.* Paper presented at the Third Phantom Users Group Workshop, M.I.T. Cambridge, MA.

Hill, E., Rieser, J., Hill, M., Hill, M., Halpin, J., & Halpin R. (1993). How persons with visual impairments explore novel spaces: Strategies of good and poor performers. *Journal of Visual Impairment and Blindness*, 295-301.

Jacobson, W. H. (1993). *The art and science of teaching orientation and mobility to persons with visual impairments*. New York, NY: American Foundation for the Blind.

Jansson, G., Fanger, J., Konig, H., & Billberger, K. (1998). Visually impaired persons' use of the PHANToM for information about texture and 3D form of virtual objects. In J. K. Salisbury & M. A. Srinivasan (Ed.) *Proceedings of the Third PHANToM Users Group Workshop*, MIT, Cambridge, MA.

Karshmer, A.I., & Bledsoe, C. (2002). *Access to mathematics by blind students - introduction to the special thematic session*. Paper presented at the International Conference on Computers Helping People with Special Needs (ICCHP), Linz, Austria.

Kitchin, R.M., & Jacobson, R.D. (1997). Techniques to Collect and Analyze the Cognitive Map Knowledge of Persons with Visual Impairment or Blindness:Issues of Validity. *Journal of Visual Impairment and Blindness*, 91 (4), 360-376.

Lahav, O., & Mioduser, D. (2000). *Multi-Sensory Virtual Environment for Supporting Blind Persons' Acquisition of Spatial Cognitive Mapping, Orientation, and Mobility Skills*. Paper presented at the 3rd International Conference on Disability, Virtual Reality and Associated Technology, Alghero, Sardinia, Italy.

Lahav, O., & Mioduser, D. (2003). A Blind Person's Cognitive Mapping of New Spaces Using a Haptic Virtual Environment. *Journal of Research in Special Education Needs, 3* (3), 172-177.

Lahav, O., & Mioduser, D. (2004). Exploration of Unknown Spaces by People who are Blind, Using a Multisensory Virtual Environment (MVE). *Journal of Special Education Technology, 19*(3).

Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.

Mioduser, D. (in press). From real virtuality in Lascaux to virtual reality today: cognitive processes with cognitive technologies.

Passini, R., & Proulx, G. (1988). Wayfinding without vision: An experiment with congenitally blind people. *Environment and Behavior, 20*, 227-252.

Schultheis, M. T., & Rizzo, A. A. (2001). The application of virtual reality technology for rehabilitation. *Rehabilitation Psychology, 46*(3), 296-311.

Semwal, S. K., & Evans-Kamp, D.L. (2000). *Virtual environments for visually impaired*. Paper presented at the 2nd International Conference on Virtual worlds, Paris, France.

Sjotrom, C., & Rassmus-Grohn, K. (1999). The sense of touch provides new computer interaction techniques for disabled people. *Technology & Disability, 10* (1), 45-52.

Sonn, U., Tornquist, K., & Svensson, E. (1999). The ADL taxonomy - from individual categorical data to ordinal categorical data. *Scandinavian Journal of occupational therapy, 6*, 11-20.

Ungar, S., Blades, M., & Spencer, S. (1996). The construction of cognitive maps by children with visual impairments. In J. Portugali (Ed.), *The construction of cognitive maps* (pp.247-273), The Netherlands: Kluwer.

Witmer, B.G., Bailey, J. H., Knerr, B. W., & Parsons, K.C. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human-Computer Studies, 45*, 413-428.

ICDVRAT 2004

Session V. Haptics

Chair: Cecilia Sik Lányi

Promoting research and clinical use of haptic feedback in virtual environments

U Feintuch[1], D Rand[2], R Kizony[3] and P L Weiss[4]

[1]School of Occupational Therapy, Hadassah-Hebrew University Medical Center, ISRAEL

[1]Caesarea Edmond Benjamin de Rothschild Institute for Interdisciplinary Applications of Computer Science, University of Haifa , ISRAEL

[2,3,4] Department of Occupational Therapy, University of Haifa, ISRAEL
[2] Beit-Rivka Geriatric Center, Petach-Tikva, ISRAEL

[1]*urif@cc.huji.ac.il*, [2]*vrand@univ.haifa.ac.il*, [3]*rachelk@zahav.net.il,* [4]*tamar@research.haifa.ac.il*

hw.haifa.ac.il/occupa/LIRT

ABSTRACT

Converging evidence demonstrates the important role played by haptic feedback in virtual reality-based rehabilitation. Unfortunately many of the available haptic systems for research and intervention are rather costly, rendering them inaccessible for use in the typical clinical facility. We present a versatile and easy-to-use software package, based on an off-the-shelf force feedback joystick. We propose that this tool may be used for a wide array of research and clinical applications. Two studies, involving different populations and different applications of the system, are presented in order to demonstrate its usability for haptic research. The first study investigates the role of haptic information in maze solving by intact individuals, while the second study tests the usability of haptic maps as a mobility aid for children who are blind.

1. INTRODUCTION

People perceive their body and the surrounding environment through different sensory modalities. Neuroanatomical and neurophysiological studies suggest that early processing is modality specific, and that later on the unimodal data are integrated into a complete description of the world (e.g., Treisman, 1986). Studies in various disciplines have demonstrated that the modalities may interact. These cross-modal interactions are observed at several processing levels. Different modalities may influence each other even prior to the appearance of stimuli, as manifested in the control of attention. Behavioral studies have shown how shifting attention in one modality caused a shift of attention in another modality. Longer lasting effects appear in cross-modal transfer, where knowledge acquired in one modality, improves performance when employing another modality (e.g., Krekling et al., 1989).

To date, most virtual environments (VE) consist of visual or audio-visual feedback. Adding haptic feedback may have several potential benefits. First, an extra channel of information may produce a more realistic environment and increase the level of presence, which consequently, may enhance the efficiency of Virtual Reality (VR)-based interventions (Durfee, 2001). Second, beyond this general enhancement which applies to any VE, specific populations may find haptic feedback especially beneficial. Many patients with stroke, for example, suffer from both motor and sensory deficits. They may benefit from haptic feedback as a component of their therapy aimed at restoring proprioceptive function. In such cases, it may be argued, that audio-visual non-haptic VR therapy may actually lead to deterioration and "learned non-use" of the affected limb. A similar case, supporting haptic feedback, can be made for interventions directed at children suffering from a high somatosensory threshold. A third, rather obvious, case where haptic feedback is required, includes populations where haptic feedback is the primary means of the intervention. Previous applications of haptic interventions include those targeted at strengthening muscles (e.g., Deutsch et al., 2001), or perceptual training for children who are blind (e.g., Colwell et al., 1998).

These concerns lead us to contend that there is a need to systematically characterize the role of haptic feedback in virtual environments and its potential benefits for virtual rehabilitation. This line of research will help developers and clinicians to decide which types of impairments (motor, cognitive, sensory) merit the

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

hardware and software costs caused by adding the haptic channel. Unfortunately, research of haptic feedback is still rather scarce. To a large extent, this void is due to the high cost and encumbrance of many of the currently available haptic devices.

This paper presents a new low-cost, user-friendly tool, aimed at facilitating multi-modal research and intervention. In order to establish the usability and justify the importance of the proposed tool, two separate studies are presented as well. We have created software which runs on a standard PC desktop and uses an off-the-shelf force feedback joystick. This user-friendly program enables a therapist or researcher to quickly design simple visuo-audio-haptic environments by drawing and encoding basic geometric shapes. Although simple to operate, many sophisticated, game-like tasks may be designed and used for a gamut of research/treatment goals that test and train participant abilities.

It should be noted that such low cost haptic devices have been suggested in the past for intervention purposes (e.g., Reinkensmeyer et al., 2002). The proposed tool, however, offers a unique and valuable feature, namely flexibility. Thus the researcher or therapist may create a wide array of environments and tasks independent of external help from a professional programmer. We believe this easy-to-use interface will make haptic research and therapy accessible to many clinicians in the “trenches”. In order to illustrate the feasibility of this tool, as well as to further establish the importance of basic haptic research, we present here two separate studies. The first study investigates the influence of haptic feedback on task performance. The second study examines the ability of haptic feedback to enhance spatial perception of children who are blind via the presentation of virtual environments.

2. THE TOOL

2.1 Hardware

We used the Sidewinder Force feedback 2 joystick manufactured by Microsoft (http://www.microsoft.com/hardware/sidewinder/FFB2.asp). It should be noted that there are other off-the-shelf products which function in similar ways (e.g., the Force 3D joystick by Logitech (http://www.logitech.com/index.cfm/products/details/IL/EN,CRID=12,CONTENTID=5016)).

2.2 Software

The software is composed of three modules. The Editor is used for designing the environments, the Simulator is used for the actual simulations, and the Analyzer serves for post-hoc analyses.

2.2.1 Editor. The Editor provides a user friendly interface (see Fig. 1) used by the investigator or therapist to place objects of different sizes and shapes on the screen. These objects are assigned various attributes such as colour, sound, movement, and type and intensity of haptic feedback. Juxtaposition of the virtual objects and association of their attributes enables the creation of either simple or complex environments. The Editor also enables the user to define the start and end points of the environment as well as the speed in which it will operate.

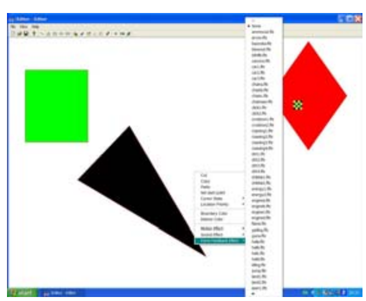


Figure 1. *Example of the Editor interface.*

2.2.2 Simulator. Once the environment is loaded to the software, The Simulator is then employed for running experimental trials or for conducting intervention. The Simulator lets the client interact with the various objects in the virtual environment. The client uses the joystick to control and move the cursor. Whenever the cursor is moved onto an object its sound and force feedback features are activated. Thus a particular sound, which may also be a pre-recorded message, will be heard. For as long as the cursor is on the object, the client will feel the haptic sensation associated with this object (as defined when creating it in the Editor).

The Simulator has several working modes suited for different client populations. The cursor can move in pre-set velocities or allow the client to change speed during the session. It can also present online the current speed and the time that has passed since the start of the session. Finally, the Simulator has a ‘right-angle’ mode where the cursor can be moved only in right angles, i.e. up-down-right-left.

2.2.3 Analyzer. The Analyzer is used by the investigator/therapist to review past sessions and analyze the data collected during these sessions. The Analyzer shows the virtual environment traversed by the client and the path of the cursor movements. The path can be viewed all at once or be animated in order to be viewed together with the client. The Analyzer may also present the temporal information associated with the trace. Thus the user may see where and at what time the cursor was at any point.

3. EXPERIMENT ONE – HAPTIC MAZE

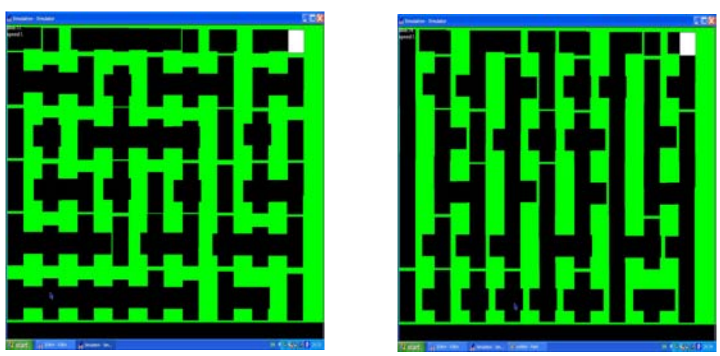
We propose this tool as a versatile system for research of haptic feedback as well as using it for intervention. Hence we decided to use it for two different experiments, each with different goal and target populations. We believe these experiments may assist to evaluate the usability of the system.

The goal of the first experiment was to test whether haptic information may facilitate performance in a maze task. This task was chosen as mazes involve both lower and higher cognitive and executive functions, and offer a variety of research and intervention paradigms (e.g., Porteus ,1973, Wann, 1997).

We hypothesized that participants may use haptic cues to help them in learning the correct route of the maze. While traversing the haptic maze, the participant constantly feels haptic feedback of many different kinds. The type of feedback is unique to each part of the maze. Our rationale was that while traversing the maze over and over, the subject will associate certain haptic feedback with the correct route, and use these haptic cues, in addition to the visual information, when solving the maze.

Our pilot studies have shown that in the case of simple mazes, the participants tended to solve them rather easily employing mainly their visual sense. They simply looked for a short while at the maze and quickly found the correct route. It seemed that, for these simple mazes, they relied almost solely on the visual information, which was so dominant, that the haptic information was not employed. These findings led us to create a dynamic maze, whose components move constantly. Parts of the maze moved back and forth either on the vertical or the horizontal axes, alternately creating and eliminating possible routes (See Figs. 2 & 3). Since these components moved back and forth repeatedly, there was, in fact, no change in the maze’s solution, and there was only one correct route. The constant motion, however, made it more difficult to detect the correct route by merely looking at the maze. This encouraged the participants to learn the maze while moving in it, thereby enhancing the opportunity to benefit from the haptic information provided to them.

Half of the subjects solved a visual-haptic maze whereas the other half solved a purely visual maze, where no force feedback cues were delivered through the joystick. We hypothesized that the first group would perform better than the latter.



Figures 2 and 3. *Screen captures of the dynamic maze at different times. The participant moves the cursor (too small to be seen in the figures) towards the white square at the top right of screen. Note how the path (black foreground) is in different positions in these two figures.*

3.1 Method

3.1.1 Subjects. Thirty six volunteers, aged 18 to 30 years (mean = 23.9; SD=3.4) participated at the experiment. All were right handed and had normal or corrected to normal vision. The number of females and males was equal. The subjects were randomly assigned to one of two groups. Each group consisted of nine males and nine females.

3.1.2 Apparatus and Stimuli. The maze was presented on a laptop computer positioned on a desk in front of the subjects, and the joystick was placed next to it. (See Fig. 4). The maze was of two possible types, visual-haptic maze or visual mazes. Otherwise they were identical in all aspects. The auditory features of the software were not employed in this task in order to limit the modalities of feedback.

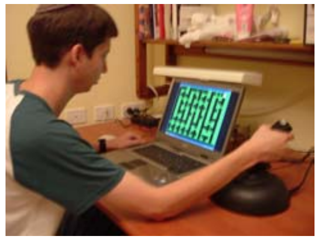


Figure 4. *Positioning of the apparatus.*

3.1.3 Procedure. Participants were first introduced to the joystick and underwent short training in its use. They were then instructed to solve the maze ten times with a short break after each trial. The execution time (ET) as well as the trace of the subject's cursor movement were collected for each trial. This session was followed by another session which took place 24-48 hours after the first one. The second session consisted also of ten trials, solving the same maze as in session one. At the end of each session the participants filled a short feedback questionnaire.

3.2 Results

The results presented here are preliminary and include only the analysis of execution times, but not the analyses of errors and of movement patterns. Generally, as commonly expected in learning tasks (e.g., Karni 1996), a learning curve was produced where execution times were improved until reaching a plateau. When comparing the mean ET of each trial of the first session to the parallel trial of the second session, a significant improvement ($p<0.05$) was achieved for all trials. The improvement of performance between sessions is known in the literature and is accounted for by the consolidation of the learned task (Karni et al., 1994). These learning effects were achieved for both groups, visual only and the visual-haptic.

To test for haptic facilitation we compared the mean ETs of the visual group to the visual haptic group. Fig. 5 describes the two learning curves (as manifested by ETs) of the first session. The differences between the groups were not significant except for the ETs of the second trial, where the visual-haptic group performed significantly faster ($t(34)=1.994$; $p<0.05$) than the visual group.

3.3 Discussion

There are two issues to be learned from this experiment. First, the data suggest that haptic feedback may facilitate learning at a rather early stage, as indicated by the faster times achieved by the haptic group at the second trial. Although this gain disappeared after subsequent trials, the learning curve took a significant 'dive' at the second trial. To further explore these findings, we plan to focus at the performance of the first few trials. This will enable us to run a much shorter experiment consisting of one short session, with many more subjects, which may lead to more decisive conclusions.

A second lesson from this experiment is the ability to use this system for research. As can be seen here, the system can be used to explore haptic feedback in motor and cognitive tasks.

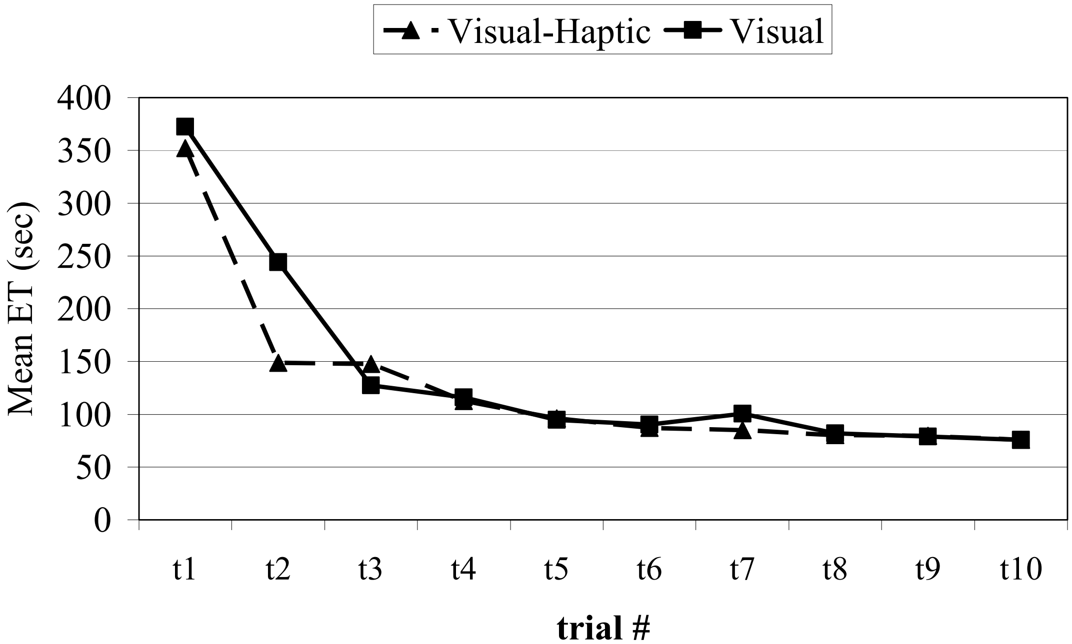


Figure 5. *Learning curves of the first session, comparing the mean ETs of the visual haptic group (dashed) and the visual group (solid).*

4. EXPERIMENT TWO – HAPTIC MAP

After using the system for a cognitive study we tested it as an intervention tool. The blind population appears to be a natural candidate who may benefit from a haptic system. Lahav and Mioduser (e.g., 2002) have used a force feedback joystick to help blind persons to get acquainted with a room and the objects within it. We took a different environment, and used our system to help children who are blind navigate in a building using a haptic map. This is an ongoing study and we report here its design as well as some very initial results.

4.1 Method

We used our system to build a haptic map of a segment of a building. In contrast to the previous experiment where the haptic feedback did not have any inherent meaning, here every type of feedback carries particular information. We created a legend of haptic feedbacks for the various classes of obstacles commonly found in buildings corridors (e.g., doors, windows, stairs, benches). Fig. 8 shows a haptic map of the areas depicted in Fig. 6 and Fig. 7. Our hypothesis is that participants who are blind may be able to use this system to learn a new environment prior to encountering it in reality.

Figures 6 and 7. *Environment to be learned by participants who are blind. The two hallways are connected and are perpendicular to each other.*

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

Figure 8. *Haptic map of the environment appearing in Figures 6-7. The different colours are used to represent different haptic feedbacks.*

4.1.1 Population. This experiment takes place in a school for the blind. Participants are children, aged 8 to 13 years, who are congenitally and totally blind. These children have not learned yet to use canes for mobility, and use primarily their hands.

4.1.2 Apparatus and Stimuli. The children use the system on a desktop in the school's computer room. The stimuli consist of a map, at a scale of about 1:100, as shown in Figure 8. The real world environment studied by the participants is located in a nearby building, where they have never been prior to this experiment. We have found two similar environments in this building so we are able to compare performance with the VR training to a baseline of no training.

4.1.3 Procedure. The participants are initially acquainted with the joystick using a very simple haptic environment. They are introduced then to the haptic map and trained with it by an instructor who verbally helps them at the early stages. As they become more proficient, the instructor gives assignments which require navigation (e.g., 'Go from the main staircase to the bathroom'). Once they perform well within the virtual environment, they are brought to the real environment and asked to perform various navigation tasks within it. They are also brought, on a separate day, to another environment in this building where they are asked to perform similar tasks. They are videotaped and their performance in both environments is compared by an evaluator, who is not aware of which environment is the learned one and which is the novel one. Each of these two environments may serve either as the practice environment or as the novel environment for different participants. The order in which these environments are introduced to the subjects is counterbalanced.

4.2 Results

As indicated above this is an ongoing study. To date we have used this system with one child (and with one adult at the early design stages) for usability testing. The child is a 12 year old girl who became blind at age 8 as a result of a tumour. She underwent a very brief training on the software, which lasted less than an hour, including the introductory session where she used the joystick for the first time ever. Upon arrival to the real environment she was asked to go to different rooms and did so very quickly and confidently, as she attested to in a subsequent interview.

4.3 Discussion

We are encouraged by this pilot result which suggests that this system may aid navigation for children who are blind. This is in line with previous studies (e.g., Lahav & Mioduser, 2002) and will likely extend their findings to the population of children and to environments of larger scale, leading to further participation of this population in everyday life.

5. CONCLUSIONS

The two studies presented here join the growing body of evidence demonstrating the important role of haptic feedback. Since many of the commercially available haptic systems require a significant investment of resources we propose this tool to help in partially filling the research void of this important area. Several studies have already successfully used force feedback joysticks for research and intervention application. This proposed tool, however, aims to be user friendly and not require the support of a programmer, thus offering both flexibility and accessibility to many researchers and therapists who wish to incorporate haptic feedback into their research

Although the tool presented here is simple to operate, many sophisticated, game-like tasks may be designed and used for a gamut of research/treatment goals that test and train participant abilities. These include:

- cognitive deficits (e.g., executive functioning, spatial orientation, attentional disorders, memory)
- motor deficits (e.g., motor planning, motor control)
- sensory deficits (e.g., orientation and navigation skills for people who are visually impaired, proprioceptive deficits for patients following stroke, re-education for peripheral nerve injuries)
- functional skills (e.g., simulator training to learn to operate a powered wheelchair)

In addition to the two example applications presented here, we anticipate that this tool will allow for future studies of cross modal tasks, eventually leading to the development of additional, haptic-based therapeutic interventions.

Acknowledgements: We thank Yaron Lachman, Barak Manos, Nili Fruchter, Carmit Gal, Orit Ziegler and Jaquline Haj for their assistance.

6. REFERENCES

C Colwell., H Petrie, D Kornbrot, A Hardwick and S Furner (1998), Haptic virtual reality for blind computer users, *Proceedings of ASSETS '98*, Los Angeles.

J Deutsch, J Latonio, G Burdea and R Boian (2001), Post-stroke rehabilitation with the Rutgers Ankle System – a case study. *Presence*, **10** ,4, pp. 416-430.

W Durfee (2001), Multi-modal virtual environments or haptics does not stand alone. *Proceedings of the Haptics, Virtual Reality, and Human Computer Interaction*, Institute, Minneapolis, Minnesota.

A Karni (1996), The acquisition of perceptual and motor skills: a memory system in the adult human cortex, *Cognitive Brain Research,* **5,** pp. 39-48.

A Karni, D Tanne, B S Rubenstein, J J Askenasy and Sagi (1994), Dependence on REM sleep of over night improvement of a perceptual skill. *Science*, **265**, pp. 679-682

S Krekling, J M Tellevik and H Nordvik (1989), Tactual learning and cross-modal transfer of an oddity problem in young children, *Journal of Experimental Child Psychology*, **47**, pp. 88-96.

O Lahav and D Mioduser (2002), Multisensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping, orientation, and mobility skills, *Proceedings of ICDVRAT 2004*, Veszprem, Hungary, pp. 213-220.

S D Porteus (1973), *Porteus Maze Test, FiftyYears' Application*, Pacific Books, Palo Alto, California.

D J Reinkensmeyer, C T Pang, J A Nessler and C C Painter (2002), Web-based telerehabilitation for the upper extremity after stroke, *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* **10**, *2*, pp. 108-102.

A M Treisman (1986), Properties, parts and object,. In *Handbook of Perception and Human Performance* (K Boff, L Kaufman & J Thomas, Eds.), Wiley, New York, pp. 1-70.

J P Wann, S K Rushton, M Smyth and D Jones (1997), Virtual environments for the rehabilitation of disorders of attention and movement, In *Virtual reality in neuro-psycho-physiology: Cognitive, clinical and methodological issues in assessment and rehabilitation. Studies in health technology and informatics* (J Riva, Ed), IOS press, Amsterdam, pp. 157-164.

Cooperative control of virtual objects using haptic teleoperation over the internet

A P Olsson[1], C R Carignan[2] and J Tang[3]

[1]Department of Electrical Engineering, Royal Institute of Technology,
Valhallavägen 79, Stockholm, SWEDEN

[2,3]Imaging Science and Information Systems Center, Georgetown University,
2115 Wisconsin Ave. NW, Suite 603, Washington, DC, USA

[1]*pontuso@kth.se*, [2]*crc32@georgetown.edu*, [3]*jt96@georgetown.edu*

www.visualization.georgetown.edu/inmotion2.htm

ABSTRACT

The feasibility of performing remote assessment and therapy of patients over the internet using robotic devices is explored. Using a force feedback device, the therapist can assess the range of motion, flexibility, strength, and spasticity of the patient's arm grasping a similar robotic device at a remote location. In addition, cooperative rehabilitation strategies can be developed whereby both the patient and therapist cooperatively perform tasks in a virtual environment. To counter the destabilizing effects of time delay in the force feedback loop, a passive wave variable architecture is used to encode velocity and force information. The control scheme is validated experimentally over the internet using a pair of InMotion2 robots located 500 miles apart.

1. INTRODUCTION

Robots have been explored as possible rehabilitation aids in laboratory settings for well over a decade. These investigations have recently expanded into the field of "teletherapy" whereby a clinician can interact with patients in remote locations using robotic devices. However, time delays encountered in the force feedback loop can cause instability in the system. Compensating for the time delay will be key to realizing this technology over the internet and is a cornerstone of the control architecture presented here.

The specific aims of our research are twofold: to enable a clinician to assess the physical condition of a patient's arm using metrics such as strength, dexterity, range of motion, and spasticity; and to help a clinician perform cooperative rehabilitative tasks with a patient using a virtual environment intended to simulate active daily living (ADL) tasks. The use of a "haptic" (force-feedback) device in conjunction with a video display will allow the clinician to remotely assess the patient's condition as well as assist the patient while performing rehabilitation tasks. The ISIS Center at Georgetown University Medical Center has recently assembled a robot rehabilitation test bed consisting of a pair of InMotion2 (IM2) robots from Interactive Motion Technology, Inc. (Krebs et al, 2001). The IM2 Robot is a direct-drive, four-bar linkage with a planar workspace of 90 x 60 cm and maximum continuous force output of 30 N in each direction (see Figure 1). The handle is pinned to the distal end of the outboard link providing a third, unactuated degree of freedom. The apparent mass at the handle is only 1.33 kg making it well-suited to our dual purposes.

This article begins with a brief review of previous work on internet therapy and cooperative haptics in Section 2. The tele-assessment and cooperative rehabilitation modes are described in Section 3. The haptic controller and time-delay compensation using wave variables are outlined in Section 4. Experimental results for both operational modes implemented on the IM2 test bed are presented in Section 5. Conclusions and future research are discussed in Section 6.

2. PREVIOUS WORK

Telemedicine has already seen several successful demonstrations of rehabilitation robotics. A "java therapy" application was enabled using a commercial, force-feedback joystick connected to an orthopaedic splint

 ISBN 07 049 11 44 2

(Reinkensmeyer, 2001). Patients log into the website "javatherapy.com" and a physical or occupational therapist will guide them through a repetitive movement regimen intended to improve their sensorimotor skills. Such therapy has been demonstrated to be useful even several years following hemiplegic stroke.

The Rutgers Haptic Master II (RMII) was used to increase hand strength in stroke patients using teletherapy (Popescu, 2000). When the patient picks up an object such as the chess piece seen on a computer screen in Figure 2, the computer actuates the piezoelectric servo valves on the hand exoskeleton to provide resistive "grasp" forces to the hand. The remote therapist can modify this resistance during the sessions to increase the patient's strength and also design an increasing complex array of virtual tasks for the patient to perform to further challenge their motor skills. Pilot clinical trials on post-stroke patients have indicated hand mechanical work increase using the RMII (Burdea, 2001).

Figure 1. *InMotion2 Robot with graphic display of virtual beam task.*

Figure 2. *Rutgers Haptic Master being used to reflect grasp force during virtual chess match.*

Cooperative control using haptic devices has been attempted on several virtual reality platforms. A pair of 2-DOF master manipulators was used to simulate thumb and index fingertip contact with an object during a peg-in-hole insertion task (Howe and Kontarinis, 1992; Burdea, 1996). Dual-arm contact with a steering wheel was simulated using a pair of 6-DOF PHANToM devices for arm motor control training (Goncharenko et al, 2003). Yano and Iwata (1995) used a pair of 6-DOF, parallel mechanism force displays to perform interactive patient-therapist tasks over the internet. Although predictive displays were used to help operators adjust for up to 3 sec delays, explicit time-delay compensation was not implemented.

Several investigators have incorporated explicit time-delay compensation in the force-feedback loops of haptic systems. Scattering theory was explored by Lawn and Hannaford (1993) to produce passive communications during teleoperation of a metal block. Wave variables were introduced by Niemeyer and Slotine (1997) for a variety of master/slave scenarios with widely varying time delay. Adams and Hannaford (1999) also considered time-delay in their passivity control formulation of stable interaction with virtual environments. However, none of the investigations we encountered considered time-delay compensation in the context of multiple haptic displays.

3. OPERATING MODES

The robot test bed has two operating modes: Tele-Assessment and Cooperative Rehabilitation. In Tele-Assessment Mode, the clinician attempts to evaluate various properties of the patient's arm through bilateral manipulation over the internet. In Cooperative Rehabilitation Mode, the patient and therapist cooperatively manipulate common objects over the internet by moving their robot handles to accomplish a therapeutic task. Both modes are described in detail below.

3.1 Tele-Assessment Mode

In this mode, the robot handle that is being grasped by the subject mirrors movements made by the clinician's robot and vice versa. A force sensor on the patient's robot relays forces exerted by the subject back to the clinician's robot where the force pattern will be "displayed" on the haptic interface. This position-based "force-reflection" is commonly used today in robot-assisted surgery.

The system block diagram for assessment mode is shown in Figure 3 and is similar to the bilateral force feedback architecture used in master/slave teleoperation. Both the master and slave are under Cartesian PD control where the position of the master becomes the desired position of the slave, and the position of the slave becomes the desired position of the master (same holds for velocity). The position and velocity data for each robot is "packetized", sent across the internet using an internet socket, and picked up by a communication process at the other side where it is unpacked and used by the local controller.

Since the PD controller filters out high frequency content from the patient's arm that might be useful for patient assessment, a force-sensor capable of picking up high frequency phenomena such as hand tremor was used to augment the haptic display. The force sensor output is high-pass filtered and transferred alongside the position/velocity data to the therapist's robot where it is amplified by a gain k and added to the PD control input. The high pass filter is necessary to remove bias readings normally present in the force sensor that would otherwise cause a position offset (Murphy, 1994).

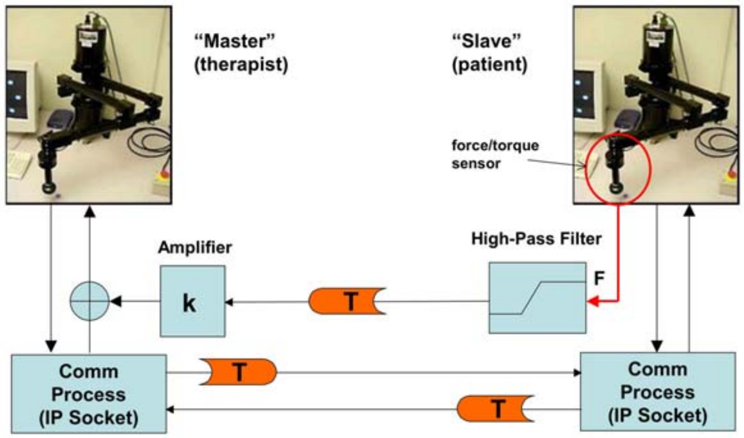


Figure 3. *Bilateral Tele-Assessment Mode Architecture.*

3.2 Cooperative Rehabilitation Mode

The control architecture for the cooperative task is shown in Figure 4. In this scenario, both the therapist and patient robots are considered "masters" which are independently interacting with the virtual object which is considered the "slave". The virtual object generator (VOG) applies the sensed "interaction" forces from the masters and then calculates the resultant motion of the object. The motion of the object at each "contact" point is then transmitted back to each master where it is tracked by a controller.

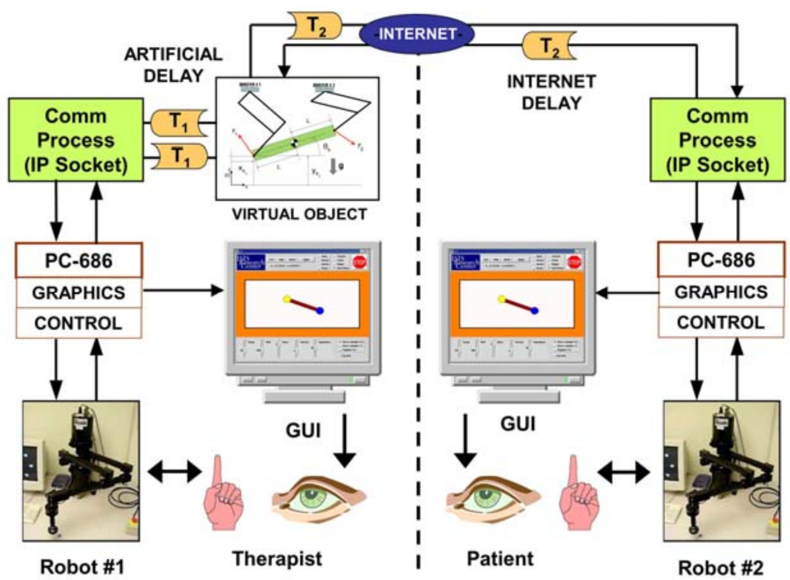


Figure 4. *Hardware configuration for Cooperative Rehabilitation Mode.*

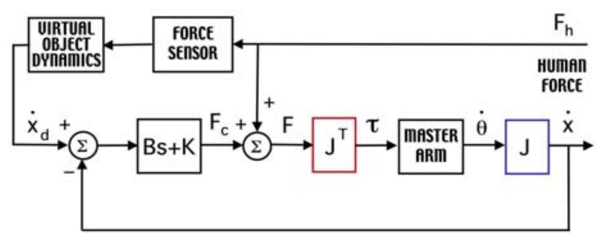
The virtual object dynamics are calculated via a separate process on one of the master arm computers. T_1 and T_2 are time delays caused by either internet transit or computational processing. If the object dynamics are being calculated on the master 1 computer, then T_1 is primarily the computational delay for the VOG process (essentially zero), and T_2 is the internet time delay for a signal to reach master 1 from master 2. However, to maintain a truly cooperative task, the two time delays should be matched. Therefore, an artificial time delay based on a moving average of the internet time delay is applied to the master control computer hosting the virtual object process (master 1 in this case).

4. HAPTIC CONTROL AND TIME DELAY COMPENSATION

In both operating modes, the core of the haptic controller is a Cartesian PD controller that servos on the position and velocity of the handle

$$F_c = B_m(\dot{x}_{md} - \dot{x}_m) + K_m(x_{md} - x_m) \tag{1}$$

where x_m is the position of the handle, F_c is the commanded Cartesian force, and B_m and K_m are diagonal damping and stiffness gains, respectively. For the cooperative mode, an additional force loop wraps around the servo loop as shown in Figure 5 to provide compliance (Carignan and Cleary, 2000). For the cooperative mode realization shown in Figure 6, the "sensed" human forces applied at each handle are used as the force inputs to the virtual object dynamics to generate the motion command inputs to each master.



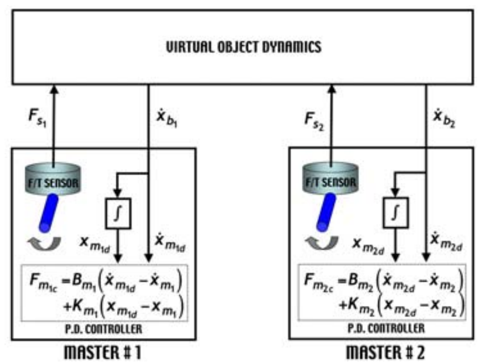


Figure 5. *Admittance controller block diagram (above).*

Figure 6. *Cooperative Rehabilitation Mode architecture using admittance control (right).*

The haptic controller works well for the interconnected robot configurations shown in Figures 3 and 4 as long as the roundtrip time delay is under about 100 msec. As the time delay starts to exceed 100 msec, the passivity of the controller becomes severely compromised and can drive the system unstable. To restore passivity in the system, compensation using wave variables emerged as the most natural approach for performing cooperative tasks over the internet (Niemeyer and Slotine, 1998).

The wave variable architecture for the cooperative mode is shown in Figure 7. The strategy is similar for tele-assessment mode except that the second master, rather than the virtual object, is the "slave". Instead of using the sensed force to impart force commands to the slave, force and velocity data are used by the master to generate an impedance "wave" command that is transmitted and decoded by the slave side into a force command for cooperative rehabilitation mode or a velocity command for tele-assessment mode. Part of the incoming wave is subsequently reflected back to the master. How much of the wave is reflected depends upon the impedance of the slave; a yielding environment will not reflect the incoming wave as greatly as a rigid wall. The wave impedance b is a tuning parameter used to trade-off speed and force; a high b produces an inertially dominant system, and a low b presents a more rigid interface (Niemeyer and Slotine, 1997b).

Each force to be "applied" to the slave is computed from the transmitted wave variable from the master using

$$F_s = -b\dot{x}_s + \sqrt{2b}\,u_s \tag{2}$$

where the incoming wave to the slave $u_s(t)$ is the delayed output wave from the master, $u_m(t\text{-}T)$. After the virtual object dynamics are computed, the virtual object generator emits its outgoing wave variable using

$$v_s = \frac{b\dot{x}_s - F_s}{\sqrt{2b}} \tag{3}$$

where the incoming wave to the master $v_m(t)$ is the delayed output wave from the slave, $v_s(t\text{-}T)$.

The desired master velocity, dx_{md}/dt, is computed from the master force F_m and return wave variable v_m as follows. The outgoing wave from the master is

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

$$u_m = \frac{b\dot{x}_{md} + F_m}{\sqrt{2b}} \tag{4}$$

If the master force command F_c is used to compute the master force in (4), then F_c=-F_m and Eq (1) and Eq (4) form a recursive loop (Niemeyer and Slotine, 1997). Substituting Eq (1) into Eq (4) and solving for dx_{md}/dt gives

$$\dot{x}_{md} = \frac{\sqrt{2b}v_m + B_m\dot{x}_m + K_m(x_m - x_{md})}{B_m + b}$$

$$x_{md} = \int_0^t \dot{x}_{md}(\tau)d\tau \tag{5}$$

Note that the wave impedance for both masters was chosen to be the same since the time delays were matched and the devices were identical.

The effect of an increase in time delay on the wave variable implementation is to decrease the system's natural frequency. The "communications stiffness" K_{comm} is given by b/T_{delay}, thus the wave impedance should be increased in proportion to the time delay to maintain system bandwidth (Niemeyer and Slotine, 1997). However, the time delay also introduces an apparent mass proportional to delay, $M_{comm}=bT_{delay}$, which produces a heavier feel at the handle as the time delay (or wave impedance) increases. Thus, a trade-off exists in wave impedance between maintaining high system bandwidth and low inertia at the handle.

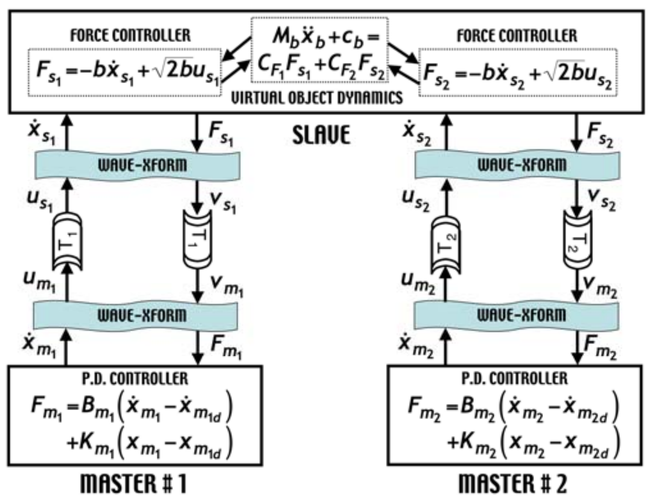


Figure 7. *Wave variable control architecture.*

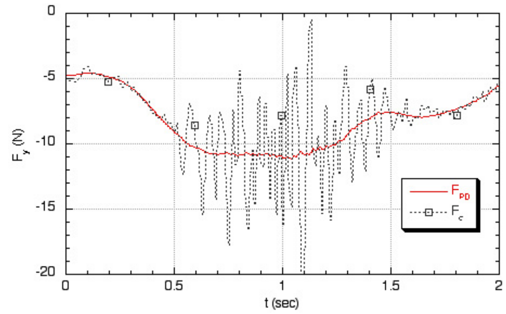
5. EXPERIMENTAL RESULTS

The control station and haptic controller operate on an AMD XP1800 PC with Athlon 686 processor running at 1533 MHz. The control process was implemented in RT-Linux and uses approximately 1.2% of the CPU time at a rate of 200 Hz. Fast internet communication between robots was achieved using UDP protocol which enabled transfer rates of 100 Hz for 16 byte datasets. A 3rd-order Butterworth filter with a 5 Hz cut-off was used for the high-pass filter in the assessment tests (Fisher, 1999). The time delay for the therapist's computer, T_1, was set equal to the internet time delay T_2 in the rehabilitation tests to maintain symmetry between the VOG and each robot.

5.1 Tele-Assessment

As a demonstration of the utility of the high bandwidth force feedback in assessment mode, an experiment was conducted in which an operator used the master robot to move the slave robot along the vertical edge of a spiral bound notebook as shown in Figure 8. The operator tried to maintain a constant normal force as the handle moved along the edge. Figure 9 shows the total force command for the master robot in the x-direction, F_c, superposed on just the PD control input force for a force gain of k=5. The ripple in F_c was due to the force sensor picking up the "tremor" caused by the spiral edge which was totally missed by the PD controller which had a bandwidth of approximately 5 Hz.

Figure 8. *Detecting the rough edge of a spiral notebook during tele-assessment test.*



.**Figure 9.** *Total force command (F_c) and PD component (F_{PD}) in x-direction for gain of k=5.*

5.2 Cooperative Rehabilitation

An example of a cooperative rehabilitation task is depicted in Figure 10. The patient and therapist "pick up" opposite ends of a virtual beam by grasping the handle which coincides with the end of the beam. Object parameters such as mass, length, and inertia can be adjusted to correspond to real-life objects using a Graphical User Interface (GUI) on the therapist's computer. The gravity vector points in the sagital plane of the operator so that s/he is pushing away when lifting the beam (toward the screen in Figure 1). As the object is "lifted", the side that is lower will begin to feel more of the weight thus stimulating the participants to maintain the beam in a horizontal position. Also, if one side tugs on the object, the other side feels it thus encouraging a cooperative strategy to lift the object.

The VOG calculates the dynamics of the virtual object being manipulated by the master arms. The centre of mass of the beam is chosen to be at the geometric centre, and the beam is assumed to be a uniform slender rod so that the inertia about its centre of mass is given by $i_b=m_bL^2/3$. The orientation of the beam with respect to the x_0-axis is given by θ_b and the total length of the beam is 2L. The resulting beam dynamics are given by

$$M_b(x_b)\ddot{x}_b + c_b(x_b,\dot{x}_b) = C_{F_1}F_1 + C_{F2}F_2 + m_b a_g \tag{6}$$

where the gravitational acceleration vector is $a_g=[0\ -g\ 0]^T$. The complete dynamics for Eq (6) can be found in (Carignan and Olsson, 2004).

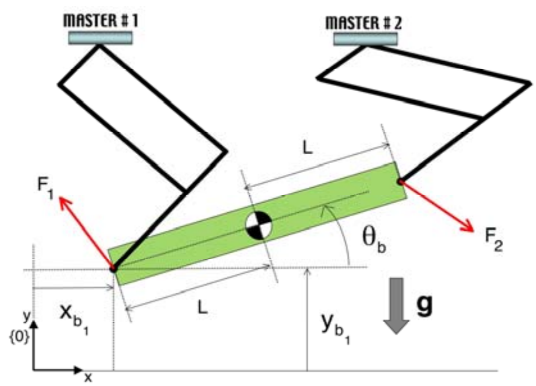


Figure 10. *Cooperative beam manipulation task.*

Three sets of experiments were performed to illustrate the cooperative beam manipulation task over the internet: admittance control with negligible delay, wave variable control for an actual internet test, and wave variable control for simulated internet roundtrip time delays of 0.5 and 1 sec. In all tests, the master controller had a bandwidth of 30 rad/sec and was critically damped yielding gains of K_m=900 N/m and B_m=60 N/m/s. The beam parameters were m_b=10 kg, L=0.15 m, and i_b=0.075 kg-m^2. A reduced gravitational acceleration of g=3 m/s^2 was used in order not to exceed the force capacity of the robot. Parameters could be changed by the operator using the graphical user interface shown previously in Figure 1.

In the first set of tests, the robots were colocated at the ISIS Center and the admittance control scheme of Figure 6 was used. The control and communication rates were 200 Hz, and the time delay within our own IP domain was only 0.15 msec. The beam starts out horizontally and then is lifted by the haptic master on the left until it reaches the vertical position. Then the second haptic master raises the right side of the beam until it is again horizontal. The plot of the beam angle θ_b versus time is shown in Figure 11.

The plots of the commanded vertical forces on the beam (sensed master forces) are shown in Figure 12. F_y for haptic master 1 is seen to go to zero when the beam reaches a vertical position while haptic master 2 sustains the full load of the beam. After master 2 raises its side of the beam, the force becomes equally distributed again. The desired versus actual velocities for master 1 (not shown) indicate very good tracking by the PD controller.

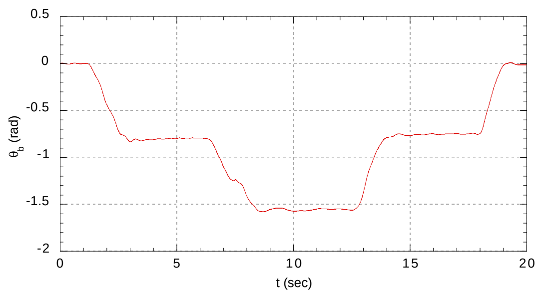


Figure 11. *Beam angle for zero time-delay test.*

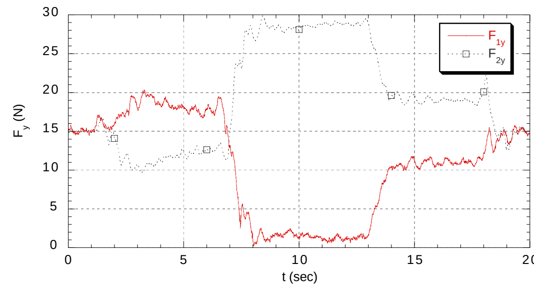


Figure 12. *Vertical force applied in zero delay test.*

In the second set of tests, the wave variable control scheme of Figure 7 was used. The controller rate was decreased to 100 Hz due to the bandwidth limitation of the communication process, and the wave impedance parameter b was set to 40 to compensate for the additional delay. The roundtrip internet time delay between Georgetown University and Cambridge, Mass. for this test varied between 35 and 110 msec and averaged about 50 msec. A 10 sec window was used to compute a moving average for the artificial delay T_1 to be applied to master 1.

The beam was manipulated in the same manner as before yielding the beam angle θ_b shown in Figure 13. The commanded master and slave forces in the y-direction for the two haptic masters are shown in Figure 14 and look remarkably similar to the zero-delay test. F_y for both masters starts out equal and then goes to zero for master 1 when the beam reaches a vertical position and master 2 sustains the full load of the beam.

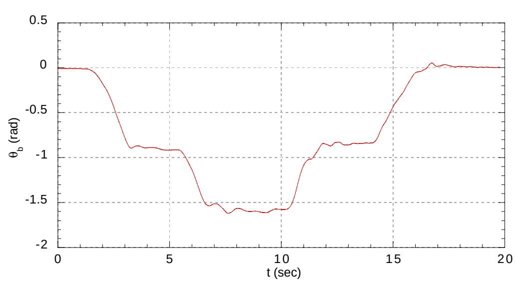


Figure 13. *Beam angle for internet test (b=40).*

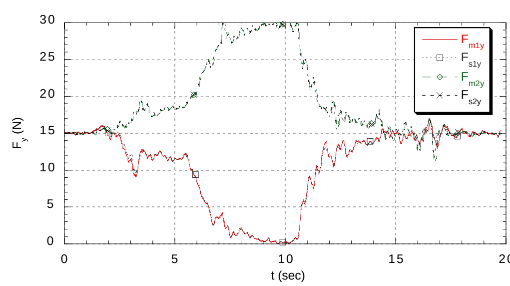


Figure 14. *Vertical force command for internet test.*

To demonstrate the feasibility of the wave variable approach for longer time delays, an internet delay simulator was used to generate 0.5 sec and 1 sec roundtrip time delays as shown in Figures 15 and 16, respectively. The decrease in system stiffness from the internet test is evidenced by the lower frequency oscillations. In addition, the apparent mass at the handle increases from approximately 0.5 kg for a 50 ms roundtrip delay to 10 kg and 20 kg for roundtrip delays of 0.5 and 1 sec, respectively. The heavier feel of the handle also made it more difficult for the operator to control contributing further to the degradation.

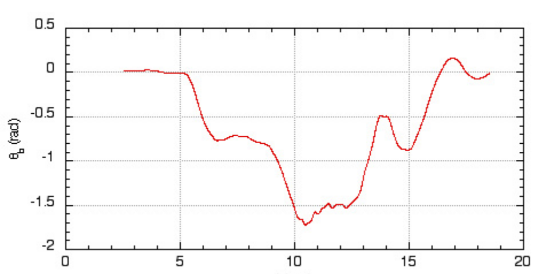
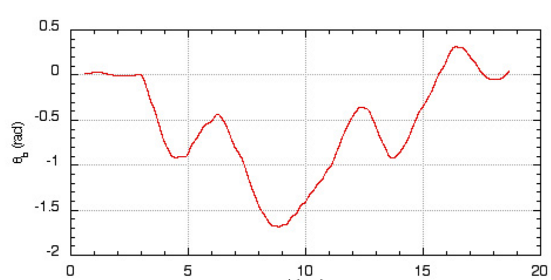


Figure 15. *Beam angle for 0.5 sec time delay test.*



.**Figure 16.** *Beam angle for 1 sec time delay test.*

6. CONCLUSIONS

Internet experiments conducted thus far indicate the feasibility of conducting both remote assessment and cooperative rehabilitation over the internet using robotic devices. During a cooperative internet task between robots 500 miles apart, time-delays of up to 110 ms produced borderline instability without compensation. However, under wave variable control, the system was robust to time-delays, and there was an almost imperceptible increase in the apparent mass of the handle. Packet loss was found to be less than 1% at transfer rates of 100 Hz when using UDP transmission.

We are currently testing even larger time-delays and examining several other cooperative tasks such as rowing and air hockey. We are also engaged in a cooperative effort with the National Rehabilitation Hospital in Washington, DC to test stroke patients over the internet and apply standard metrics for assessment. In addition, a head-mounted display and tracker are being integrated into the system to allow for more realistic simulations using 3D visualization. Coordination of the haptic and visual feedback in the simulator (stereopsis) is an area of ongoing research as are strategies for dealing with packet loss during less reliable transmission.

Acknowledgements: This project is being supported by the U.S. Medical Research and Material Command under Grant #DAMD17-99-1-9022.

7. REFERENCES

R Adams and B Hannaford (1999), Stable haptic interaction with virtual environments, *IEEE Trans. on Robotics and Automation*, 15(3), pp 465-474.

G Burdea (1996), *Force And Touch Feedback For Virtual Reality*, John Wiley and Sons, New York.

C Carignan and K Cleary (2000), Closed-Loop Force Control for Haptic Simulation of Virtual Environments, *The Electronic Journal of Haptics Research* (http://www.haptics-e.org), 2(2), pp. 1-14.

C Carignan and P. Olsson (2004), Cooperative Control of Virtual Objects over the Internet using Force-Reflecting Master Arms, *Proc. IEEE Int. Conf. on Robotics and Automation*, New Orleans, pp. 1221-1226.

J Craig (1989), *Introduction to Robotics: Mechanics and Control, 2nd ed.*, Addison-Wesley, Reading, Mass.

T. Fisher (1999), Interactive Digital Filter Design, http://www-users.cs.york.ac.uk/~fisher/mkfilter/

I Goncharenko, M Svinin, S Matsumoto, S Hosoe and Y.Kanou (2003), Design and implementation of rehabilitation haptic simulators, *Proc. Intl. Workshop on Virtual Rehabilitation*, Rutgers, pp. 33-39.

R Howe and D Kontarinis (1992), Task performance with a dextrous teleoperated hand system, *Proc. of SPIE*, 1833: pp. 199-207

H I Krebs, N Hogan, W Hening, S Adamovich and H Poizner (2001), Procedural motor learning in parkinsons disease, *Experimental Brain Research*, 141: pp. 425-437.

C Lawn and B Hannaford (1993), Performance testing of passive communication and control in teleoperation with time delay, *Proc. IEEE Intl. Conf. on Robotics and Automation*, Atlanta, pp. 776 –783.

S Murphy and D Robertson (1994), Construction of a high-pass digital filter from a low-pass digital filter. *Journal of Applied Biomechanics*. 10: 374-381

G Niemeyer and J-J Slotine (1997), Designing force reflecting teleoperators with large time delays to appear as virtual tools, *Proc. IEEE Intl. Conf. on Robotics and Automation*, Albuquerque, pp. 2212-2218.

G Niemeyer and J-J Slotine (1997b), Using Wave Variables for System Analysis and Robot Control, *Proc. IEEE Intl. Conf. on Robotics and Automation*, Albuquerque, pp. 1619-1625.

G Niemeyer and J-J Slotine (1998), Towards force reflecting teleoperation over the internet, *Proc. IEEE Intl. Conf. on Robotics and Automation*, Leuven, pp. 1909-1915.

V Popescu, G Burdea, M Bouzit, and V Hentz (2000), A Virtual-Reality-Based Telerehabilitation System with Force Feedback. *IEEE Trans. on Information Technology in Biomedicine*, 4(1), pp. 45-51.

D Reinkensmeyer (2001), Java Therapy: A Web-Based System for Mass-Delivered Movement Rehabilitation After Stroke, *Proc. of the State of the Science Conf. On Telerehabilitation,* M. Rosen and D. Lauderdale, eds., pp. 70-73, Washington, D.C.

H Yano and H Iwata (1995), Cooperative work in virtual environment with force feedback, *Proc. 7th Intl. Conf. on Artifical Reality and Tele-existence (ICAT'97)*, pp. 203-210

Providing external memory aids in haptic visualisations for blind computer users

S A Wall[1] and S Brewster[2]

Glasgow Interactive Systems Group, Department of Computing Science, University of Glasgow,
17 Lilybank Gardens, Glasgow, G12 8QQ, UK

[1]steven@dcs.gla.ac.uk, [2]stephen@dcs.gla.ac.uk

www.multivis.org

ABSTRACT

Haptic force feedback devices can be used to allow visually impaired computer users to explore visualisations of numerical data using their sense of touch. However, exploration can often be time consuming and laborious due to the "point interaction" nature of most force feedback devices, which constrains interaction to the tip of a probe used to explore the haptic virtual environment. When exploring large or complex visualisations, this can place considerable demands on short term memory usage. In this respect, a fundamental problem faced by blind users is that there is no way to mark points of interest or to harness external memory, in a similar way in which a sighted person may mark a graph or table at a point of interest, or leave a note in a margin. This paper describes the design, implementation and evaluation of external memory aids for exploring haptic graphs. The memory aids are "beacons" which can be used to mark, and subsequently return to, a point of interest on the graph. Qualitative evaluation by visually impaired users showed that external memory aids are a potentially useful tool. The most commonly reported problem was that of using the keyboard to control placing of the beacons. Suggestions for subsequent re-design of the beacons in light of the participants' comments are considered.

1. INTRODUCTION

Understanding and manipulating information using visualisations such as graphs, tables, bar charts and 3-dimensional (3D) plots is a very common task for sighted people. The skills needed are learned early in school and then used throughout life, for example, in analysing information, creating presentations to show to others, or for managing home finances. The basic skills needed for creating and manipulating graphs are necessary for all parts of education and employment. Blind people have very restricted access to information presented in these visual ways. It is currently very hard for them to create, manipulate and communicate visualisations such as graphs and tables. As Wies *et al.* state "Inaccessibility of instructional materials, media, and technologies used in science, engineering, and mathematics education severely restricts the ability of students with little or no sight to excel in these disciplines. Curricular barriers deny the world access to this pool of potential talent, and limit individuals' freedom to pursue technical careers" (Wies et al., 2001). Traditional methods of presenting visualisations to blind and visually impaired people include Braille diagrams, heat-raised paper, screen readers and screen magnifiers. There are several drawbacks inherent with these methods in that they are either unable to respond quickly to dynamic changes in data (hard copies need to be produced of Braille and heat raised diagrams which is often slow and difficult for a blind person without sighted assistance), they are inherently serial in nature and therefore highly memory intensive (e.g. a screen reader reading values from a graph or a Braille table), use an abridged form of the data (pre-recorded descriptions of graphs delivered via spoken word or Braille versions of tables), or are simply inaccessible to potential users (only 26% of visually impaired university students read Braille, and screen magnifiers are useless to those with no residual vision).

It is increasingly important to provide fast and reliable access for visually impaired people to the proliferation of digitally stored data, including that which is available on the internet. The EPSRC-funded Multivis project is a collaboration between the Departments of Computing Science and Psychology at the University of Glasgow investigating tools to allow visually impaired people access to data visualisations.

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

Techniques from virtual reality are used to present the data using multiple modalities, in particular, haptics and audio. Haptic force feedback devices and tactile displays potentially provide a richer method of interacting with digitally stored data than those currently available to blind persons. Using a haptic device, a blind person could edit and perceive data in real time, whilst working alongside sighted colleagues. Many of these devices have been designed with the desktop in mind (for example, the desktop PHANToM from Sensable Technologies). Some mouse type devices are small and discrete enough to pass as standard computer mice (the Wingman force-feedback mouse from Logitech, or the Virtouch VTPlayer mouse). Work on the Multivis project has extensively employed the PHANToM force feedback haptic interface (See Figure 1). It consists of a kinematic framework with three rotational degrees of freedom, allowing for exploration of a 3D Cartesian workspace (13x18x25 cm). The user interacts with the device by gripping a stylus attached to the distal point of the framework. The device is nominally passive (it does not resist the motion of the user), but motors located on each of the joints can be selectively activated to convey the illusion of contact with a rigid surface.

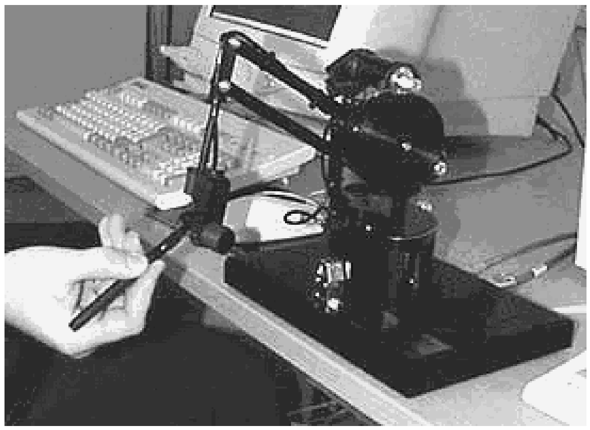


Figure 1. *The PHANToM haptic interface from Sensable. The user grips the stylus to interact with the device. Movement of the framework is selectively constrained to portray the illusion of contact.*

Previous research has shown that in the absence of visual information, users are able to perceive and interpret multimodal (haptic and audio) representations of common graph types such as line graphs, bar charts and pie charts. Experimental results showed that a multimodal representation of line graphs was significantly more accurate than a raised paper based representation, however, exploration times were significantly slower (Yu et al., 2002). This increase in time was attributed to the point interaction nature of the PHANToM. Limiting the user to a single point of contact precludes the use of Exploratory Procedures (Lederman and Klatzky, 1987) such as enclosure and contour following that are important for perceiving size and shape of objects efficiently (an essential action for comprehending the data in graphs). The lack of spatially distributed cutaneous information on the finger tip means that the users instead have to integrate a series of temporally varying cues as they traverse the graph. Exploration is therefore slow and highly memory intensive as little context can be provided through a single point of stimulation. These problems are further exacerbated when dealing with large data sets or data exhibiting a high dimensionality.

A fundamental problem faced by blind people when interacting with visualisations (or any complex information) is that there is no easy way to mark points of interest or to access external memory (Zhang and Norman, 1994); a sighted user might mark a graph with a pen to indicate an interesting point to return to later, or write something in the margin as a reminder. Such external memory is a very powerful tool for sighted people and can significantly reduce working memory requirements. This is not possible for blind people and means that they may easily get lost in the data, overloaded, and makes it hard for them to mark interesting points in the data. This slows down interaction, increases workload and means that it is more likely that mistakes will be made. As Stevens suggests, providing access to an external memory aid will give very substantial benefits to blind users (Stevens, 1996). This paper describes the design and evaluation of an external memory aid for blind and visually impaired users accessing complex visualisations using a PHANToM force feedback device.

2. INTITIAL DESIGNS

This section describes the design of the external memory aids that are proposed. Also provided is a brief description of the haptic graph rendering software, which is necessary to give context to the evaluations in Sections 3 and 4. Readers are referred to the papers referenced herein for a more complete description of the system.

2.1 External Memory Aids

The external memory aids were initially designed to be used with three dimensional surface plots of data rendered using the haptic interface. The PHANToM force feedback device could be used to explore the height and contours of a surface which represented a three dimensional data set, for example, as illustrated in Figure 2. Data were stored in a table for different combinations of x and y values. The height of the surface on the z axis is proportional to the value of the data for that combination of x and y. The large surface area of the plot could potentially present spatial memory problems for a visually impaired person exploring through the single point of contact offered by the haptic device.

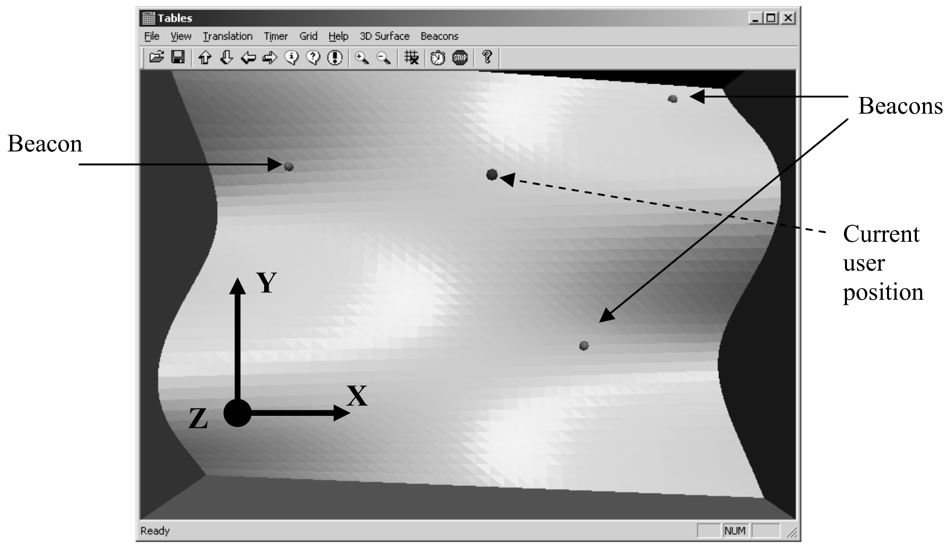


Figure 2. *Screen snapshot of a haptic surface plot with visual representation of beacons. The user can feel the 3D surface of the table using a PHANToM haptic interface. The value of each cell is mapped to height in the z- direction.*

The initial design of an external memory aid consisted of multimodal beacons using haptic and audio cues. By positioning the PHANToM at a point of interest on the 3D plot and issuing the appropriate command (via a key press) a beacon could be enabled at that point (see Figure 2). There were three beacons available to the user; one beacon was assigned to each of the keys “a”, “s” and “d”, and was enabled by pressing the relevant key. Users could subsequently return to a beacon they had placed by pressing the relevant key again, along with the “shift” key, to access a “seek” mode for that beacon. This caused the PHANToM to guide users to the beacon’s location, by actively dragging them from their current position to the position of the beacon, using a virtual spring force. The keys used to control the beacons were chosen so that they could be used with the non-dominant hand, while grasping the PHANToM stylus with the dominant hand. Thus, having identified a salient point on the graph, for example, a local minimum, should the user find a second local minimum, he/she can easily compare the two without devoting time to relocating the original point. Non-speech audio MIDI percussion sounds were used to represent the beacons. The beacons were differentiated by each using a different timbre. The audio was panned to the right or left relative to the current position of the PHANToM end point. The volume decreased exponentially with distance from the beacon. This helped provide some context regarding the relationship between the user’s current point and any points of interest he/she may have marked.

2.2 Haptic Bar Charts

As the haptic surface plots have not yet been formally evaluated with visually impaired users, we opted to perform evaluations of the beacons using existing software for rendering haptic bar charts. These have been tested several times with visually impaired users and have been shown to be a robust design (Yu and Brewster, 2003). As the main purpose of this investigation was to assess the design of the external memory aids, using them with the untested surface plots may have produced experimental difficulties that were attributed to the memory aids themselves, but which were actually flaws in the design of the surface plots. The size of the bar charts was increased from seven bars, used in previous studies, to twelve bars, to increase the associated memory demands.

The virtual bar charts used were rendered using the GHOST SDK from Sensable Technologies as used in previous studies of the Multivis project (Yu and Brewster, 2003). The bars are located on the back wall of the workspace facing the user, as opposed to on the “floor” of the virtual environment, as with traditional raised paper graphs on a desk or table. A snapshot of a graphical representation of this environment is illustrated in Figure 3. The bars are constructed out of polygons that form a V-shaped cross-section. The purpose of the V-shaped channel is to retain the PHANToM pointer within the line. Preliminary studies with haptic line graphs showed that users had problems keeping the pointer on raised objects (Yu et al., 2001). A concave shape is an effective solution to this problem. The user could click the PHANToM stylus switch while in a bar to have the label for the bar read out (a text label describing the bar provided in the data file). No information was given on the data value of the bar. There were 12 bars in each chart; the user could thus feel the height of the bars using the PHANToM stylus in order to make a comparison of the heights.

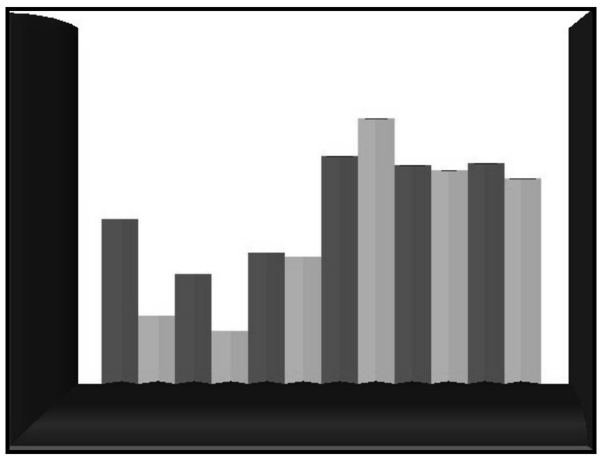


Figure 3. *Visual representation of a haptic bar chart with 12 bars, as used in the evaluation. Bars were rendered as grooves in the surface in order to constrain the user’s exploration.*

3. PILOT STUDY

Prior to testing the beacons, an informal pilot evaluation was undertaken with two participants from the University of Glasgow. This study was conducted to evaluate the basic design of the beacons before they were presented to the visually impaired users, and to correct any obvious design flaws. It is important to maximise the productivity of evaluations conducted with visually impaired users; pilot testing with sighted users (in the absence of visual information) is an efficient method of refining the designs of stimuli and experimental procedures. Neither participant was naive to the purpose of the study, and both had full (or corrected) vision. The participants were not blindfolded but were unable to see the monitor displaying the visual representation of the graphs. Both participants were presented with several bar charts and asked some questions on the data to encourage them to use the beacons (the full experimental procedure used in the formal experiment is outlined in Section 4).

Two significant design modifications occurred as a result of the participants’ comments. Firstly, it was noted that the users placed the external memory aids less often than they used the “seek” function. The general strategy employed was to place the beacons on the bars of interest and then subsequently use them to jump between the bars and compare the heights. Participants also became frustrated when they moved a beacon by accidentally pressing the “place” button for the beacon more than once. As placing beacons was a less frequent action, the controls were changed so that a modifier key press was needed to place the beacon, and seeking was the default action (enabled by pressing the relevant beacon key alone). This also reduced the chances of accidentally replacing one of the beacons with an erroneous key press. Secondly, the participants found the audio cues from the beacons confusing and distracting to use. As such, the recurrent audio cues were disabled, and instead a percussion noise was played once, to let the users know they had successfully placed a beacon. This revised design of the external memory aids was then evaluated with visually impaired participants.

4. EXPERIMENTAL PROCEDURE

To evaluate the external memory aids, a user-centred experimental design was employed to capture blind and visually impaired users' requirements and opinions. The main purpose of the study was to identify areas in which the design of the beacons could be improved using qualitative data obtained via a *post hoc* interview with the visually impaired participants. To stimulate opinions on the beacons, participants were given tasks to perform in a condition with the beacons, and in a condition where they did not have access to them. The views of the participants could then be used within an iterative design process to create a second, refined version of the beacons.

Eight participants took part in the evaluation. All were registered blind and based at the Royal National College for the Blind in Hereford, UK. They were all paid for their participation in the evaluation. When verbally questioned during briefing on the experimental procedure, all participants said they had prior experience with the concept of bar charts before participating, except for one participant. This participant was shown an example of a tactile raised paper bar chart and had important elements such as bar and axes described verbally by the researcher. The participants were free to ask any questions and gave verbal consent when they felt they had grasped the concepts sufficiently to progress with the evaluation. The participants also had varying degrees of residual vision. Three of the subjects had used the PHANToM before on previous experiments run as part of the project. All subjects were given full instructions on how to operate the PHANToM and given a verbal walkthrough of the elements of a haptic bar chart (as described in Section 2.2) by the researcher.

The two conditions of the study were with and without the external memory aids. The order of these conditions was counterbalanced between participants to control for ordering effects in the data. For the condition incorporating the memory aids, the participants were encouraged to use them to aid their performance. A group of fifteen bar charts were used in the experiment, based on data gathered from National Statistics Online (http://www.statistics.gov.uk/) and WorldClimate (http://www.worldclimate.com/). The same group of fifteen bar charts was used in both conditions of the experiment to prevent any confound due to the relative difficulty of different sets of graphs. To prevent the participants from using prior knowledge obtained during the first condition to aid performance in the second condition, they were told that the graphs represented different information. Thus, in the first condition, the participants were told that the graphs represented monthly coffee export figures for different countries. An example graph from this set is a bar chart showing the average monthly coffee exports for Haiti, with the bars labelled from January to December, or the average coffee exports for the month of June over 12 different countries, the bars being labelled with the country names in alphabetical order from left to right. In the second condition the data were described as weather statistics such as rainfall, temperature and pressure levels from cities around the world. An example bar chart from this set might be the average monthly rainfall for Glasgow, with the bars representing the different months, or the average temperature over 12 cities for a particular month, with the bars labelled with the names of the cities.

There were 15 trials during each of the experimental conditions. In each trial the participant was presented with one of the bar charts in the data set. Prior to its presentation, the participant was told what the bar chart represented (for example, "Coffee exports for Burundi over the months of the year") and given a question to answer as quickly as possible using the information in the bar chart. This question was always of the form of which bar had the highest/lowest value out of three given bars (for example, "During which month were the coffee exports for Burundi highest, out of May, June and October?"). All the information needed to answer the question was contained within the graph. There were no trick questions, and participants were informed that they were not required to answer any further questions regarding the graph and therefore need not concern themselves with any other bars, provided they locate the three salient to the question posed. However, participants were advised that they should check the bars thoroughly, rather than assume the answer from any knowledge they had on the subject of the question. Participants had a 2 minute time limit to answer the question, with a speech audio reminder provided by the software after 1 minute. Participants verbally indicated their answer to the researcher, who immediately stopped the software timer and made a note of the response. The participant was then presented with the next trial in the condition, using the same procedure. The proportion of correct responses and the average time taken to answer each question was used as a measure of performance for the participants in each condition. It was hypothesised that during the condition in which the beacons were used, the participants would perform faster, as they would be able to use the beacons to return quickly to bars salient to the question, in order to make comparisons more swiftly.

A post-experiment interview between the researcher and participant was conducted, during which the participants were invited to give their opinions on the system, and the researcher also questioned them regarding any strategies or behaviours they had employed in use of the external memory aids. All interviews were recorded with participant's consent and later transcribed for analysis.

5. RESULTS

Although the participants were encouraged to use the beacons by the researcher during the relevant condition, ultimately, it was at the participants' discretion as to how much they employed the external memory aids. The frequency with which participants used the beacons varied widely, from regularly to very infrequently. This can be attributed to several factors which varied between individual participants. These were: level of aptitude with the haptic device, degree of residual vision, level of expertise in using the keyboard, and additional impairments related to loss of vision which the researchers had to be sensitive to (e.g. motion impairment in the non-dominant hand used to operate the keyboard). This was potentially compounded by the fact that the beacons were not particularly easy for the participants to use, as their comments revealed during the interviews. Therefore, it was decided that a formal, quantitative analysis of beacon use would not give the most useful data to help refine the design, and analysis therefore focused on the qualitative comments from the interviews.

Averaged over both conditions and across all participants, the proportion of correct responses was 76% (standard deviation = 13) and the average time taken to answer a question was 50.92s (standard deviation = 12.48 s). This illustrates that the participants had little problem answering the questions accurately, and within the time limit set by the researchers, nor was the task so easy that a ceiling of performance level was reached.

From analysis of the transcribed user comments obtained during the interviews, the general feeling of the participants was that the external memory aids were a potentially useful addition to the haptic graphs. The majority of the participants stated they would find the memory aids useful in certain situations, even if they did not employ them extensively during this study. The most common scenarios for use of the memory aids were to avoid searching erroneous bars while traversing large distances between different extremes of the graph:

> *"When two are very similar … it does help to go straight to it, because it's distracting when you look at the height of a country that you don't need."*

> *"If you need to get back, like you're at one end of the track … this thing seems a million miles apart to me because I'm blind, so if you're up in December and you want to go straight to January … in the end I let it go and it went back and found it"*

> *"It was easier when you were moving greater distances, easier with the markers if you were going from one end of the graph to the other, it was more convenient to help you find your place."*

Two of the participants, both of whom had previously participated in experiments using haptic devices (but without the external memory aids), felt that they did not need the external memory aids to perform the task. They both expressed the view that the memory aids would be more useful to beginners finding their way around the graph. Despite this, both concurred independently that the memory aids would potentially be more useful with more complex questions or graphs:

> *"I wouldn't say it would never be useful. It'd be nice to know it was there … if I was looking for information on two things like what are the two highest months of the rainfall."*

> *"I'd say if I had large, four, five or six countries to find, marking the important ones for me might be interesting."*

The most common negative remarks were regarding the use of the keyboard to place the beacons, and the demands on memory associated with use of the beacons. Four of the participants commented directly on the problems associated with using the keyboard at the same time as the PHANToM. Participants found it distracting to use the keyboard whilst attempting to maintain a steady position with the PHANToM stylus on the bar they were exploring. This often led to beacons being placed on bars other than those intended, which was confusing for participants. A salient point was that, despite having the external memory aids introduced to them as "beacons", only two participants explicitly referred to them using this term in the *post hoc* interviews. A further participant referred to using "the keyboard" and only used the term "beacons" after being prompted by the researcher. The remaining participants referred to them as: "keys", "keyboard" or "buttons" (3 participants); "markers" or "marking" (2 participants); no explicit reference (1 participant).

Coupled with the generally low usage of the beacons, this possibly indicates that the participants had difficulty grasping the concept of "beacons" as presented during the study. This may be due to the fact that in other contexts, "beacons" are often auditory or visual in nature, whereas the "beacons" used in this study were neither. This would suggest that a change in name would be useful for further investigations.

Several of the participants indicated that they felt they were able to cope with the memory demands of the questions without recourse to the external memory aids (3 participants). Two of the participants disliked the extra memory demands that were incurred through actually using the beacons:

> *"It would be handy if it told you where you've already placed it …at the moment you're having to use your memory a lot … and then you still have to concentrate on the heights of the rainfalls or whatever."*

> *"Another feature would be to bring up a list of beacons that you've already marked because I sometimes forgot what I'd marked."*

These memory demands seem qualitatively different from participants remembering where bars were spatially, in relation to one another on the graph. This did not seem to pose a problem for most participants; several commented on the fact:

> *" … the second and third time I knew the order of the countries really quickly … If I was to use the beacons a lot, I think that would've held me up more than just letting me get on with using my memory … I could just remember where Haiti was."*

> *"I thought it was interesting how I would just jump to the three countries mentioned or the choices mentioned without knowing the rest of the graph, it was almost like I knew where it was and I've got no sight at all ... I could visualise the graph … I do visualise stuff quite a bit, and that seems to have played a part …"*

One participant likened the memorising of locations of the bars to remembering positions of favourite tracks on a new album:

> *"I soon quickly remembered them, but that made me think of when I play a CD at home and it's the first time … I listen to the whole album all the way through … I remember the tracks as it's running through … I can remember it, and I've only listened to the CD once. That's the same as with countries … I only had to look through that once and I could more or less remember the order …"*

Two participants both suggested that a speech reminder as to where the beacons were placed should be available by pressing an easily accessible button, the space bar, for example.

There were several other comments raised by one or two individual participants that are potentially interesting for influencing a second design iteration of the external memory aids. Two participants commented on the audio feedback. One participant felt that the percussion noise when placing a beacon was not audible enough, and speech output (e.g. "beacon activated") would be better, particularly in a noisy environment such as a classroom or a communal office. The second participant suggested that a different warning noise should be played if the user attempted to place a beacon that had already been placed earlier. With regards to the haptic interface device, one participant indicated a preference for being able to use their fingers to explore the graph rather than the stylus. Similarly, another participant suggested that Braille output for the bar labels might be useful if incorporated in to the PHANToM:

> *"What would be really nifty, is incorporating a small Braille display in to the pen. So as you press the button* [on the stylus] *it pops up under your fingers, "Beijing", or whatever…"*

6. DISCUSSION

Given the participants reported experiences with the external memory aids, it is evident that there are several areas where the design needs to be improved in order to make the potential benefits more accessible. The most significant problem appears to be in the use of the keyboard, simultaneously with the PHANToM stylus. In particular, the use of the arbitrary keys assigned to the beacons, and the use of the modifier key which necessitated two fingered interaction appeared to cause problems. Several participants implicated this as the reason that they chose not to use the beacons regularly, or the reason they found that progress slowed down when they were used. Another potential reason for the low level of beacon use is that the participants found the concept of the "beacons" difficult to grasp. This is demonstrated by the fact that not many participants referred to the beacons directly during the *post hoc* interviews without prompting from the researcher. Several participants chose to describe them in terms of the "buttons" or the "keys". The participants chose to characterise the memory aids by the physical interface, rather than its function within the system, which suggests they may not have fully grasped the potential of the beacons and developed strategies for their use. Despite this, the participants almost unanimously (7 out of 8 participants) thought the external memory aids were potentially useful, in particular for novice users, more complex data, or for traversing large data sets quickly.

An improved design of the beacons should therefore attempt to improve the accessibility of the keyboard interface, leverage the visually impaired users' excellent spatial memory skills, and be conceptually simpler, to suggest a potential model of use to participants. The proposed solution at present is to allow the participants to "snap to" any bar in the chart by pressing a corresponding number on the numeric keypad. Thus, there is no "setting" of beacons, freeing the participants of this aspect of the cognitive load. Most visually impaired users are comfortable with the numeric keypad as it is often employed in screen reading software (for example, JAWS from Freedom Scientific). Provided they remember the relative position of the bar in the graph (something most participants excelled at) they could press the corresponding numeric key to move to the bar. In this way, the system is made more analogous to using a TV or CD player remote control in order to skip to tracks/channels/bars of the user's choice; a concept which most of the users should be familiar with. This would eliminate erroneous placing of memory aids, and the need to remember their positions, whilst allowing the participants to use their spatial memory skills to remember the location of the bars. This method of interaction is less generic and could not be applied to the haptic surface plots without redesign, but could potentially solve many of the problems encountered with bar charts.

7. CONCLUSIONS AND FUTURE WORK

This paper has considered a qualitative evaluation of a preliminary design of external memory aids to be used in haptic bar charts for visually impaired users. The users seemed to appreciate the potential of the memory aids, but their comments revealed several shortcomings of the current design. Participants' comments were used to suggest a further iteration of the design which will attempt to reduce the cognitive load of using the memory aids whilst providing the benefits highlighted by the participants. Future work includes implementing the new design with both bar charts, and more complex 3D surface plots, as originally envisaged. It is also planned to perform a more longitudinal study with the memory aids in order to identify common actions, recurrent problems, potential shortcuts and emergent strategies for use.

Acknowledgements: This work was supported by the EPSRC grant GR/S53251/01, "An investigation of the use of tactile displays for visualisation for blind people". We would also like to thank all at the Royal National College for the Blind at Hereford for their hospitality, access to the students, and for allowing the evaluations to be run at the college.

8. REFERENCES

Lederman, S. J. and Klatzky, R. L. (1987). Hand movements: A window in to haptic object recognition, *Cognitive Psychology* **19** 3,pp. 342-368.

Stevens, R. (1996). Principles for the design of auditory interfaces to present complex information to blind people, Ph.D. Thesis, University of York, UK.

Wies, E., Gardner, J., O'Modhrain, S. and Bulatov, V. (2001). Web-based touch display for accessible science education. In *Haptic human-computer interaction*. (Brewster, S. A. and Murray-Smith, R. Eds.). Berlin, Springer-LNCS, **2058**, pp. 52-60.

Yu, W. and Brewster, S. A. (2003). Evaluation of multimodal graphs for blind people, *Journal of Universal Access in the Information Society* **2** 2,pp. 105-124.

Yu, W., Ramloll, R. and Brewster, S. (2001). Haptic graphs for blind computer users. In *Haptic human-computer interaction*. (Brewster, S. A. and Murray-Smith, R. Eds.). Berlin, Springer LNCS, **2058**, pp. 41-51.

Yu, W., Reid, D. and Brewster, S. A. (2002). Multimodal virtual reality versus printed medium in visualization for blind people. *In Proceedings of ACM ASSETS*, Edinburgh, Scotland, pp. 57-64.

Zhang, J. and Norman, D. A. (1994). Representations in distributed cognitive tasks, *Cognitive Science*, pp. 87-122.

Developing a multimodal web application

A Caffrey[1] and R McCrindle[2]

[1]Department of Cybernetics, University of Reading,
Whiteknights, Reading, RG6 6AY, UK

[2]Department of Computer Science, University of Reading,
Whiteknights, Reading, RG6 6AY, UK

aidancaffrey@hotmail.com, *r.j.mccrindle@reading.ac.uk*

[1]*www.geocities.com/aidancaffrey,* [2]*www.cs.rdg.ac.uk*

ABSTRACT

This paper describes the creation of a multi-modal website that incorporates both haptics and speech recognition. The purpose of the work is to provide a new and improved method of internet navigation for visually impaired users. The rationale for implementing haptic devices and speech recognition software within websites is described, together with the benefits that can accrue from using them in combination. A test site has been developed which demonstrates, to visually impaired users, several different types of web application that could make use of these technologies. It has also been demonstrated that websites incorporating haptics and speech recognition can still adhere to standard usability guidelines such as Bobby. Several tests were devised and undertaken to gauge the effectiveness of the completed web site. The data obtained has been analysed and provides strong evidence that haptics and speech recognition can improve internet navigation for visually impaired users.

1. INTRODUCTION

Many inventions have had major effects on people's lives – changing the way they perceive the world around them. Perhaps the most recent example is the internet; since the development of HTML by Tim Berners-Lee in 1990 the usage of the internet has grown exponentially, with corresponding effects on peoples lives. A large percentage of internet traffic contains digital images and with the increase of internet speed, linked to the uptake of broadband connections the growth in transmission of digital images will continue. With more visually impaired users needing to access web products with high levels of graphic imagery, there is a clear need to address the difficulty they may encounter in order to ensure that they are not disadvantaged in a work or social context.

To date, the main methods of web navigation for visually impaired users have been screen readers or a Braille display. However, both these technologies currently fail to address many of the problems faced by visually impaired internet users. For example, web navigation by these technologies is frequently slow and often some of the information on a web page remains unavailable to them. There are a large number of visually impaired people worldwide - over two million in the UK alone. Thus, there is a great demand for improved technology to assist visually impaired internet users with web navigation and web based tasks.

There are several areas of research underway into how to improve internet usage for visually impaired people. One of the most prominent is speech recognition; originally this technology was excellent when computers only had a textual display and a small set of commands to deal with. However, as computer interfaces and applications increased in complexity, the speech models failed to keep up with this rapid pace of development and consequently speech recognition systems do not provide the same degree of internet usage satisfaction to visually impaired users as is available to sighted users browsing via more traditional means.

Another very new way to represent information is through haptics. Information can be conveyed to a user through the sense of touch. This technology in general, appears to have substantial potential (as shown by Wall and Brewster, 2003; Yu et al, 2003b), however integrating haptics into web pages is a novel concept and requires further research into its potential. Lahav and Mioduser (2002) showed that combining haptics and speech recognition could give a better overall transmission of information than either on its own. However caution must be applied to this idea - Gladstone et al (2002) showed that excessive or conflicting information (via haptics and speech recognition) can reduce the information transfer process.

A generic problem associated with being visually impaired is the inability of the user to fully mentally map spaces (Lahav and Mioduser, 2002). From this it follows that visually impaired users may not be able to map movements of a mouse to curser movements on a screen – causing problems with targeting (e.g. hyperlinks). Keates et al (2000) invented *gravity wells* – which are a novel method of helping motion-impaired users with targeting tasks.

The main devices for input to a computer are still keyboard and mouse. There are other (haptic) devices that can be used such as tactile mice, trackballs, Gamepads, joysticks and wheels. All these devices have slightly different properties and are good at carrying out different tasks; to try to improve web navigation it makes sense to use a mouse. Currently there is an excellent device available, the Logitech Wingman Force Feedback mouse (shown in Figure.1), that offers full force feedback (which unlike some other devices means it can actually move the device rather than just impede its movement in a direction) and only costs around £50 – so it is affordable by the majority of users.

Figure 1. *Logitech Wingman Force Feedback mouse.*

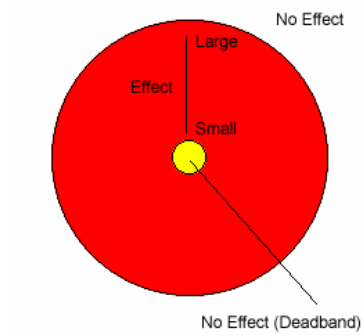


Figure 2. *Where haptic effects are felt in and around an image*.

The novel aspects of the work described in this paper is the implementation of the gravity well idea such that it become much easier for visually impaired people to use a mouse. The particular application chosen to demonstrate this is web navigation – thereby solving two problems simultaneously. Additionally, we use speech recognition and haptics in combination, in order to compensate or circumvent any weaknesses inherent in each technology. The web site created shows several applications of how speech recognition and haptics may be of use to visually impaired users.

2. IMPLEMENTATION OF HAPTIC EFFECTS

2.1 Gravity Wells

The idea of gravity wells originates from the effect found around black holes, whereby all matter around a black hole becomes drawn to its singularity (centre). In computer terms: the curser is the matter, the black hole is an object on the screen and the singularity is the centre of the object. The curser is drawn to the centre only when it has entered the object.

In their original implementation, gravity wells assisted motion-impaired users with their targeting of tasks by applying a haptic effect to move a device to the centre of the target when the curser reached a certain proximity to that target. One of our main tasks is to adapt this approach so that it can be useful for visually impaired users.

2.2 Creating the Gravity Well Effect

The chosen shape for a gravity well is circular, so that the effects experienced are consistent regardless of where the curser enters the image of a gravity well. This representation however is inconsistent with the frequently encountered rectangular representation of images. This creates a mapping problem. To feel effects for only a circle portion of that image, an ellipse mask may be used. This masks the part of the image in which effects should be experienced from the rest of the image - when the mask is entered or exited, effects can be started or stopped respectively as per general gravity well implementation.

Either a spring or a slope may be used to simulate the gravity well effect. Firstly, considering the spring, when the curser is at the edge of the gravity well, the spring is at full extension, so the pull to the centre (and the springs resting position) is the greatest. In and around the centre of the gravity well is a dead band where no effects are felt - this is to allow the curser (and device) to settle in the middle and avoid problems of oscillation and over-usage of motors. The strength of the spring is proportional to the distance from the centre, up to the edge of the gravity well where the strength is the maximum possible. Figure.2 illustrates these concepts.

Secondly, a slope can be used in the place of a spring. The effects are similar, however, instead of the spring extending, the slopes steepness increases as the edge of the gravity well is approached. Again there is a small dead band around the centre. The slope can be customised more precisely than the spring – so has the potential to provide the better effect; unfortunately it is not yet supported by the Immersion Web Plug-in and hence the spring was selected for use in this project.

2.3 Coding the Effects

There are several programs provided with the Wingman mouse such as Immersion Studio 4.1.1 which can be used to choose effects, and their parameters and store them in Immersion Force Resource (IFR) files. However, due to differences in Macromedia Dreamweaver MX version to previous versions, using these files is not yet feasible. This means that the haptic effects need to be coded in HTML sending commands to the Immersion Web Plug-in directly. All the effects need a function to create and initialise them, a function that calculates the position of the centre of the gravity well and starts the effect, and a stopping function.

2.4 Haptic Problems and Solutions

The main problems faced were as follows: there is a lack of primitives that can be represented well as effects, the haptic interfaces cannot simulate the effects accurately enough and sometimes haptic effects can confuse and unnecessarily complicate situations – making the transfer of information worse.

The first problem is dealt with by selecting primitives (e.g. spring) that are well represented as effects. The Wingman mouse is an excellent haptic device and can simulate the required effects well thus solving the second problem. The third problem is combated by carefully selecting when to use effects – keeping them to a minimum and not conflicting with other outputs.

3. IMPLEMENTATION OF SPEECH RECOGNITION

3.1 System Setup

For useful speech recognition the following is needed: text to be read out from a web page, constrained commands to be received and processed from a user, ability to navigate between web pages, and text-to-speech to inform the user what will happen when clicking gravity wells.

The speech engine used is the Microsoft English (U.S) v6.0 Recogniser - this allows speech recognition and text-to-speech. Also needed to allow speech recognition to work in web pages is the Microsoft Internet Explorer (I.E.) Speech Add-In. Finally a standard microphone and set of speakers are required.

3.2 Voice Web Studio

Only since 2002, when Speech Application Language Tags (SALT) was released, has speech recognition in web pages been feasible. SALT extends other languages to provide speech capabilities, in this case the language extended was HTML. An Extension to Dreamweaver MX called Voice Web Studio (VWS) was used to include SALT into HTML. VWS offers four main types of speech elements: speech prompts, speech listens, dialogs and play dialogs.

‘Speech prompts’ are text-to-speech where either text on the screen is read out or pre-defined hidden text is read out. ‘Speech listens’ are elements that listen for user speech input, then compare the input to a list of pre-defined commands and process the result. ‘Dialogs’ are speech prompt and speech listens running in conjunction, to form a simple dialog between the user and the computer. For example a speech prompt may ask the user a question, to which they answer and then receive another speech prompt in response. Finally ‘Play dialogs’ are special speech elements that can be bound to objects in a web page, such as images. They can be used to play a speech prompt to the user in response to events including OnMouseOver.

To make speech recognition work in a web page that page must first be ‘speech activated.’ This means selecting which prompt/listen/dialog will play when the page loads and choosing the browser type. It is

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

standard in this work for a dialog to start whenever a page is loaded to identify the page and allow the user to instantly navigate around the page/site.

The speech navigation model used in the web site is able to access, at any point, a set of standard commands and at some points be able to access some specialist commands. The standard commands will allow the user to move to any page within the site, access help or turn off speech. The specialist commands will only work on certain pages and wherever possible their use will not block subsequent use of the standard commands. This is achieved by predominantly using dialogs so speech listens are active most of the time, allowing fast and easy speech navigation for new and experienced users.

3.3 Speech recognition Problems

Using speech recognition demands a substantial CPU resource. This can cause older machines to lag, especially when the CPU is performing other tasks, listen elements can appear to not work under conditions of lag. This problem can be solved, most of the time by not running unnecessary applications and using a highly specified PC. This problem is likely to persist however, as newer and more complex speech models are continually being developed to utilise the growing speed capacity of newer PCs.

Another major problem is mis-recognitions. This is when valid commands are not recognised. This can be caused by varying accents, user age and sex. There is no standard user voice; the software used is designed for a US voice – the people tested are all from the UK. This can be reasonably dealt with by training the speech recognition software to a users' voice; VWS had event handlers for events such as mis-recognitions that can limit the impact of the problem.

4. WEB PAGE CREATION

4.1 Page Layout

The web pages have been constructed with consistency in mind. The difficult mental mapping task need only be carried out once (so when a user has learnt to use one page, they should be able to use the entire site). As per standard there is a navigation frame on the left of the page (which will contain the gravity wells); the information frame is on the right. Users will always enter into a dialog when a page loads and will start a play dialog whenever they enter a gravity well (giving advice on what clicking the gravity well will do).

The gravity wells are laid out in a vertical stack of three (in the navigation bar) with spacing between them to allow easy distinction between them. There are three because then the most common user requests in a web site are covered: loading the site index, loading the next page and re-loading the current page. Also this design is extendable to enable more layers to be added to the site – instead of loading the site index, a sub-index could be loaded and so forth.

Finally there is an animated gif composed of a series of slightly differing fractals as the background of the navigation bar. The fractals were made in Ultrafractal 3.0.3 and bound together in Macromedia Fireworks MX. The last frame is similar to the first to allow for a smooth changeover. The main reason for this is so that the navigation bar is very striking from the rest of the page – thus making the gravity wells easier to locate by partially sighted users. A additional reason is to make the page(s) aesthetically pleasing to sighted for users, for reasons of recommendations etc.

4.2 Gravity Well Design

The gravity wells are designed so that they can be easily located – this is done by keeping them in a consistent place on the page and increasing their size sufficiently. Also they have been designed such that the users need not necessarily click in the centre of the gravity well, any point inside it will suffice, thereby reducing the time needed for navigation.

Gravity wells also have a visual element to them; when the curser is not in one they have a very bold design (helps partially sighted users with targeting their location). When the curser enters the gravity well a short animated gif plays; it visually mimics the feeling of being drawn to the middle of the gravity well by the haptic interface. This adds to the immersion for sighted and partially sighted users. The last frame is very different to the first, this sudden 'jump' between frames keeps attention on them.

4.3 Pages Created

Several different types of web pages were created during this project in order to demonstrate the different potential areas of use of this technology to visually impaired people. Initially when the site loads, a page that is almost text only (so it will work fully with previous technologies) details the software/hardware needed to fully operate the site. Also the choice of a normal version or a plain background version of the site is offered. Also for greater accessibility purposes a small collection of text-only pages containing explanations of the features in the pages was created.

The first main page is the site index, this page introduces the site (and the project behind it), offering a more detailed background if required. The index page is shown in Figure. 4. From this point onwards speech and gravity well navigation is enabled. The second page is the training page, this offers the user an opportunity to understand how to use speech recognition and find and use gravity wells. This is a very important page because ideally it removes the need for prior training and users can teach themselves how to use the site.

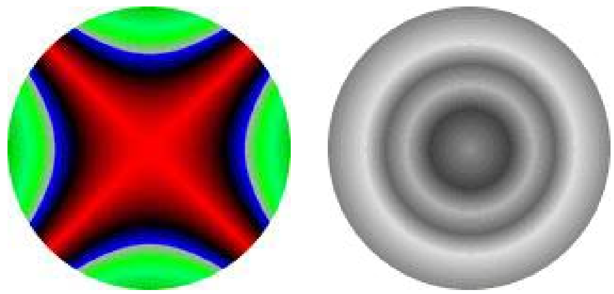


Figure 3. *A Gravity well: when (left) the curser is not in it and (right) when the cursor enters the well.*

Figure 4. *Index page of gravity web.*

The third page is news page, on this page users may choose articles to be read out and skip between paragraphs on the articles. This page is designed to replicate and improve on the function of a screen reader – showing off the new speech software now available for use on the internet.

The fourth page has links to two animations that contain haptic effects within them. This first of these is a flash animation of a picture moving around the screen, the Wingman mouse follows the movement of the picture. The second is of a flash button, haptic effects are felt when clicking the button. Both the animations were made in Macromedia Flash MX and are designed to show the potential for haptic effects in animations.

The fifth page is a comments page, on this page text-fields can be filled via voice commands and then emailed to the webmaster. This page is designed to mimic the process of filling in forms online, such as online shopping or emailing and show it can be done by speech recognition. The sixth page is a publications page where relevant works can be found. The purpose of this page is to demonstrate the potential for loading other websites within the sites frame structure and also to demonstrate how it would be possible to linked back from these external sites to increase awareness and publicity of the site.

4.4 Achieving Accessibility

For a project such as this achieving some form of accessibility was very relevant. Two different free online testing services were used: Bobby and UsableNet – Bobby was the service predominantly used. A URL can be passed to Bobby and a list of accessibility errors are generated in response. Many of the errors were easily solvable such as stating the language of the page, providing alternative content for images and ensuring that fonts are set to relative rather than exact sizes. Solving these problems allows previous technologies such has screen readers to use the pages better.

There are different levels of accessibility award available from Bobby; the standard is Section 508 approval and the highest is Bobby AAA approval. The vast proportion of the site we developed achieved Bobby AAA approval and the entire site achieved at least Section 508 approval.

As haptics and speech recognition on the internet are both very new it is not surprising that errors were generated. For example, one of the main problems found is the use of the 'object' tag (which is essential) because this tag is not supported by older browsers or PDAs. Of course older browsers do not support haptics or speech recognition anyway, but nevertheless descriptive tags of what would happen if the tag was supported are needed.

Another accessibility issue considered was MarkUp Validation. The problem with HTML browsers is that where code is not clear, the browser will try to guess what was meant. Not only may the guess be wrong, it may also be different from browser to browser. The validator used is the WC3 MarkUp Validation Service; as with Bobby a URL can be given and a list of errors is generated. The same problem occurred – as the software used the newest tags available old versions of HTML were shown not to be able to support them. However the process generated a number of errors that were solvable, thus further increasing the sites accessibility.

5. TESTING

5.1 Gravity Well Testing

The original implementation of gravity wells (Keates et al, 2000) was shown to aid targeting times. Our implementation has changed so much from this original application that a short test to determine whether the new implementation is also effective at aiding targeting was devised. The test involved asking a set of users to move the curser to a corner of the screen and then move the curser into one of the gravity wells. This was then carried out for each of the remaining gravity wells and then all repeated from each of the corners. The test is carried out twice by each user, once with haptic effects and once without - so the times can be compared. All the usual steps to ensure fair testing such as half the users starting with haptics and the other half without were taken.

Over the twelve targeting tasks carried out by each user, every user experienced a saving in time with haptic effects, with the average time saved per target just over 0.2s and in total 2.5s. The slowest user took 17s to complete the task (without haptics), making a saving of 15%, up to the fastest user of 10.5s, making a saving of 24%. Unsurprisingly the time taken to reach a target was related to the distance to it, barring a few anomalies. Figure 5 shows the cumulative targeting times for all the users.

5.2 Web Site Testing

In this test speech recognition and haptics are working together, this test is important because it replicates the conditions users exploring the site on the internet would experience. The purpose of these tests were to gauge the effectiveness of incorporating speech recognition in the site (commands were tested three times each), the ability of users to find gravity wells with increasing experience and a series of qualitative questions to obtain information about the effectiveness of the site in general and possible improvements to all aspects of it. Yu and Brewster (2002) suggest that there is no significant difference in test performances between blindfolded users' test results and blind users' test results. Working on this assumption, our tests were carried out with a group of students from The University of Reading, for some parts of the tests they were blindfolded to simulate blindness.

Looking first at the speech recognition results, 86% of commands were correctly recognised. No significant difference between the genders was found. Apart from a few instances all users were able to get any command recognised at least one out of three times. The highest percentage of commands correctly received from a user was 93% moving reasonably linearly to the worst of 79%.

When users were requested to find the gravity wells blindfolded, all users experienced a reduction in the overall time needed to complete the test from their first attempt to their third. The average time to complete a single targeting task fell from 3s at the first attempt to 2.4s at the third (20% reduction). Most users improved at each attempt. The total time all users needed dropped 11% from the first to the second and a further 7% from the second to the third (shown in Figure 6). Also notable was that users experienced a drop in the number of times they entered an incorrect gravity well.

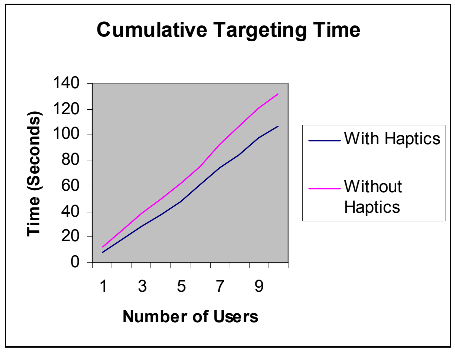


Figure 5. *Cumulative targeting times for users with haptics (bottom line) and without (top line).*

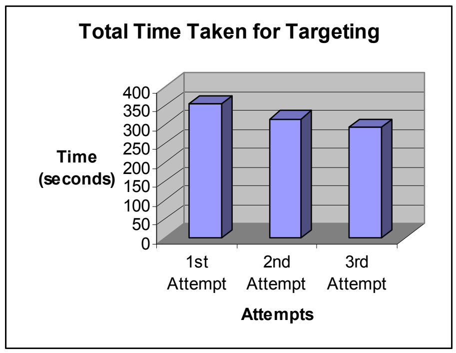


Figure 6. *Total times taken for targeting by all the users at each attempt.*

Finally some of the qualitative information obtained revealed: a general dislike of the computer generated voice, some words were not being pronounced correctly by the speech engine, haptic effects were very popular – especially in animations, the gravity wells were also popular and have a reassuring effect on users.

5.3 Analysis and Explanation

The results obtained provide encouraging evidence that not only the method in general but also the exact implementation of gravity well synchronised with speech recognition is a viable and promising method of web navigation. Further, the evidence demonstrates that this is the case for all users – as to varying extents all users need to carry out targeting and navigating tasks.

The speech recognition test showed that all users received high recognition percentages. Caution must be exercised however, as a high percentage is often still not acceptable enough. However, the perceived standard of speech recognition technology by users in general, is often not that high and this view may have contributed to the percentages we achieved being seen as favourable by the users taking part. It should be noted that extensive testing was carried out prior to the user test and any commands that proved hard to recognise were replaced if it was possible to do so. When a command was not recognised it was usually on the first attempt made by the user, on failing they would then speak more clearly and more often than not the command was then recognised. On a few occasions users were simply not able to pronounce a command in such a matter that it would be recognised, so failed on all three attempts made.

There were a few anomalies within the gravity well testing (where haptics appeared to increase targeting time) that can be explained by observation made during the experiment. User targeting was sometimes a little off, leading to them being caught momentarily in the gravity of an adjacent gravity well. Sometimes users moved the Wingman mouse very quickly and moved through the gravity well and out the other side and then had to move back into it.

As with most tasks the users were able to learn how to find the gravity wells and perform it faster. Mental mapping and muscle memory may well be the cause of this. Backing up that claim is the fact that users curser trajectory was shaky and non-linear at the first attempt, becoming a more direct and linear movement at the second and third attempt. Once the trajectory was learned the users increased the speed at which they moved the curser.

6. CONCLUSIONS

This project investigated the feasibility of a new implementation of gravity wells, with new speech recognition software, for improved web navigation for the visually impaired. The results show that this approach has a high feasibility to provide significant improvements on how visually impaired people navigation the web and hence improve their quality of life.

The new speech software performed well, with good event handlers to deal with problems on the occasions when it did fail. The haptic effects developed have shown that haptics can improve the internet for all users, particularly in animations. Several effects were developed and the one selected was received very well.

The Wingman mouse is an excellent device that performed well for this project, simulating the desired effects well. Some users complained of slight discomfort when using it, despite this it should be considered for any haptic project where the complexity of effects needed is not too great.

Performing accessibility and MarkUp checks was very important not only for this project, but for all web sites in general. The problems encountered were due to the checks requiring that legacy software is able to run the HTML code. When using the newest software available this is not always possible to fully achieve; the process of trying to achieve accessibility is very worthwhile as accessibility will increase even if the highest standards are not quite met.

Within this area of research there is a lack of projects being carried out. There are a vast number of disabled (in this case visually impaired) users – so there is a large demand for more extensive research into ways of providing equal access to the internet for these users. Replacing the information lost by not having a visual channel is a difficult task that haptics and speech recognition cannot yet achieve fully. However there is great promise in the future of these technologies and further research is required.

Having successfully proved our initial concepts, it is our intention to carry out further testing and development with the involvement of target groups of visually impaired users.

Acknowledgements: We would like to thank the students at The University of Reading who took part in the testing for free.

7. REFERENCES

K Gladstone, H Graupp and C Avizzano (2002), Assessing the utility of dual finger haptic interaction with 3D virtual environments for blind people, Proc. 4th Intl. Conf. Disability, Virtual Reality and Assoc. Tech, Hungary.

S Keates, P Langdon, J Clarkson and P Robinson (2000), Investigating the use of force feedback for motion-impaired users, CNR-IROE, Florence.

O Lahav and D Mioduser (2002), Multisensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping, orientation and mobility skills, Proc. 4th Intl. Conf. Disability, Virtual Reality and Assoc. Tech, Hungary.

S Wall and S Brewster (2003), Assessing haptic properties for data representation.

W Yu and S Brewster (2002), Comparing two haptic interfaces for multimodal graph rendering, Proc. 10th Symp. On Haptic Interfaces for Virtual Environments and Teleoperator Systems, USA, Florida, pp. 3-9.

W Yu, K Kangas and S Brewster (2003), Web-based haptic applications for blind people to create virtual graphs, 11th Haptic Symposium, Los Angeles.

ICDVRAT 2004

Session VI. Audio Virtual Environments

Chair: Tomohiro Kuroda

Design and user evaluation of a spatial audio system for blind users

S H Kurniawan[1], A Sporka [2], V Nemec[2] and P Slavik[2]

[1]Department of Computation, UMIST
PO Box 88, Manchester M60 1QD, UK

[2]Department of Computer Science, Czech Technical University in Prague
Faculty of Electrical Engineering, Karlovo náměstí 13, Praha 2, 12135, CZECH REPUBLIC

s.kurniawan@co.umist.ac.uk, *sporkaa@fel.cvut.cz, nemec@fel.cvut.cz, slavik@fel.cvut.cz*

ABSTRACT

The paper reports on the design and evaluation of a spatial audio system that models the acoustic response of a closed environment with varying sizes and textures. To test the fit of the algorithms used, the system was evaluated by nine blind computer users in a controlled experiment using seven distinct sounds in three environments. The statistical analysis reveals that there was insignificant difference in user perception of room sizes between sounds in real and simulated scenes. This system can contribute to the area of VR systems used for training blind people to navigate in real environments.

1. INTRODUCTION

Virtual Reality (VR) has been an important and exciting field for many years. Recently, its potential for people with disabilities has picked up. VR systems have been applied in the areas of education, training, rehabilitation, communication and information technology for people with disabilities (Colwell, et al., 1998).

One important application of VR is to train blind users to navigate and move around in real environment, also known as the orientation and mobility (O&M) training (Inman and Loge, 1999). O&M training is important because it helps blind people develop the skills and techniques to overcome travel difficulties created by blindness and to maximise their ability to move around in different environments independently, safely and confidently (The Royal Blind School, 2003).

Conventional O&M training involves exposing trainees to various environments to train them to detect the sound variation caused by environmental factors, e.g., the floor textures, the room size, the location of the closest obstacle, etc., or instructing a blind trainee to approach a wall or an obstacle to show the sound variation caused by the presence of object, known as the obstacle perception training (Seki and Ito, 2003). This method is very time consuming and in may pose some danger to the trainees (e.g., when training them to cross a busy road). This is an area where VR may be beneficial. However, this also means that the acoustic system used for the training must be able to simulate the sound variation caused by the environments.

This paper reports on the design and evaluation of one component of the O&M training system for blind and visually impaired people: a spatial audio system that is capable of modelling the acoustic response of a closed environment with varying sizes and textures (e.g., a small carpeted room vs. a large marble hallway).

2. SPATIAL SOUNDS IN THE REAL AND VIRTUAL ENVIRONMENTS

2.1 Spatial Sound

Sound is a vibration of particles around their equilibrium positions. The vibration of the particles causes small local adiabatic variations of pressure in the medium, referred to as the *acoustic pressure*. Through the environment, these variations are propagated by means of waves of acoustic pressure.

In the real world, containing obstacles among the sound sources and receivers, only some of the sound wave may travel directly between the source and the receiver (*the direct sound*). Signal 1 in Figure 1 is an example of a direct sound. Except for the change of its intensity due to energy dissipation, and temporal

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

displacement due to the finite phase velocity of the sound waves, the shape of the signal of a direct sound is unchanged. Other parts of the original sound energy will be reflected or diffracted by the obstacles of the environment before reaching the receiver.

Combined from all contributions, the receiver obtains the acoustic response of the environment to the sound emitted from its source. As described later, the acoustic response may be understood as a compound of the early echoes and late reverberation. Figure 1 illustrates sound propagation in a closed environment. The acoustic response of Figure 1 can be represented as a diagram with the time of arrival (at the receiver) of all echoes on the X-axis and the intensity of sound on the Y-axis, known as the Impulse Response (IR) diagram.

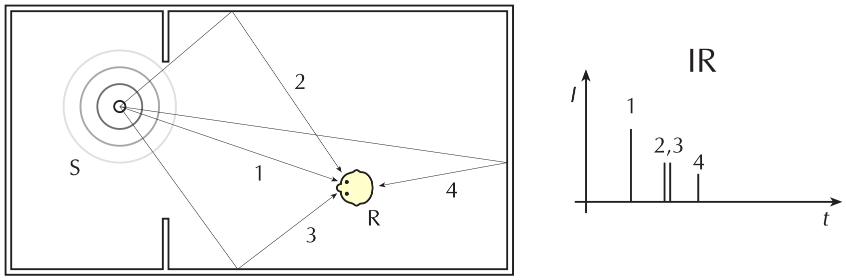


Figure 1. *Propagation of sound in a closed environment. S = sound source, R = sound receiver. The diagram on the right hand side is the corresponding IR diagram.*

2.2 The Human Auditory System

Because the human auditory system is capable of the detection of the reverberation in the sound received and the analysis of the spatial description of the surrounding environment contained in it, it is necessary to include the acoustic response of the environment to make the rendered sounds in the VR systems sound natural to their users. As the sounds interact with different obstacles while propagating, the sounds that arrive to the listener contain multiple echoes of the original signal emanated from the source (a combination of sounds with varying delays and magnitudes of attenuation), and also the information about the directions of arrival.

The echoes may be divided into the following groups:

- The first echo received by the listener is interpreted by their auditory system as *the direct sound*. Its direction of arrival gives the most significant information about the position of the source. Its intensity gives a hint on the distance of the source.
- *The early echoes* (within the first 100 ms) allow the listeners to tell the position of the nearest obstacles from the analysis of their incoming direction and intensity (Funkhouser, et al., 2002).
- *The late reverberation* carries the information about the overall layout of the environment (the size of the room, the textures of the floor/wall, etc.)
- The human receives the information about the direction of the incoming sound by analyzing the difference of the sound signals received by each ear. The following effects are of special interest:
- Interaural Time Difference. As sound travels at relatively low speed and human ears are spatially separated, there is a slight time difference between the sound arriving to the left ear and to the right ear.
- Interaural Intensity Difference. Because of its mass, human head becomes an obstacle for the ear at the opposite side of the sound source.
- Pinnae Response. Due to their high geometric complexity, ear auricles (pinnae) act as very sophisticated filters whose characteristics are highly dependent on the direction of the incoming sound. The analysis of attenuations of signal in certain frequencies allows human auditory system to determine the vertical component of the direction of incoming sound signal.

All these phenomena are modelled as the Head-Related Transfer Function (HRTF), a response function describing the acoustic filter of the environment near a listener's head. The HRTF depends on the position of a sound source relative to a listener and the size and shape of the torso, head and pinnae (Algazi, et al., 2002).

2.3 Spatial Sound in Virtual Environment

A sound propagation in any environment may be described using an acoustic wave equation (Kuttruff, 2000). However, its solution is very complex even for simple configurations and virtually impossible for more complicated scenes where many obstacles are involved. Therefore, some alternative ways to describe sound waves are necessary. We may distinguish three different types of approaches to the approximation of the wave equation solution: numerical, geometrical, and statistical.

2.3.1. The numerical approaches. These approaches provide an approximation of the wave equation by reducing the problem to estimation of energy transfers among finite elements specified within the modelled scene. There are the following two principal approaches:

1. *The finite and boundary element methods* provide a solution to the wave equation by dividing the space of the modelled environment into distinct elements. The continuous wave equation is converted into a discrete set of linear equations. The underlying computation of these methods is generally very complex. As the necessary resolution of the spatial subdivision increases with the highest frequency of the sound to be modelled, these methods are suitable only for simulation of low-frequency energy transfers within simple scenes.
2. *The waveguide mesh* is a regular array of elements with its neighbours connected by unit delays. Each element describes the sound energy of a finite part of the modelled environment. Each sound source and receiver is modelled as an element from the mesh either with input or output of the signal. The simulation itself is iterative. During each iteration, every element updates its energy status based on the previous energy status of all its neighbours following the energy conservation law. The IR is then described by the development of sound energy in the receiver (Lokki, et al., 2002).

2.3.2 Geometrical approach. In these approaches it is assumed that the sound wavelengths are smaller than the obstacles in the scene and therefore they are usable only for the simulation of sounds of high frequencies. However, their smaller computational costs compensate for their inaccuracy and make them usable in various VR systems. The paradigm of these approaches is the sound wave simulation through the investigation of the sound rays. The audibility of sound sources in the position of the listener is determined and quantified through finding the rays that represent audible echoes of the emitted sound. The following two methods should be mentioned:

1. *Ray tracing*, adopted from the well-known method in the 3D computer graphics, is based on the concept of sound rays tracing as shown in Figure 2. Each ray of the initial set of rays emanating from the sound source *S* is traced and compared with the position of sound receivers *R1* and *R2*. The tracing is stopped when a certain condition is met (e.g., the maximum order of reflection or the minimum level of energy has been exceeded, or the ray hits the receiver).

2. *Beam tracing* is based on the concept of tracing the sound beams. A beam is a cone defined by its apex (sound source) and its base (a closed environment). as illustrated in Figure 3. A beam represents all rays that would originate in the beam's apex and intersect the beam's base. Using this method, larger areas of the space are searched at once as illustrated in Figure 4.

3. The beam tracing is a generalization of the ray tracing method. The result of the tracing process of a single beam is a beam tree in which each beam is represented by a node. The children of these nodes represent the beams that originate from the collisions of their parent beam with the obstacles. Calculating the reflections of a beam is more computationally expensive than calculating the reflections. Since the number of beams increases exponentially in reflections of higher order, this method is only usable for early echoes. Consequently, it is impossible to use this method to simulate the long reverberations, as the reflections of high order (e.g., ≥20) are not computed in reasonable time.

Formally, the beam tracing may be described as (Heckbert and Hanrahan, 1984):

```
algorithm beamtracing
input: scene in boundary representation
output: beam tree
method
    create an initial set of beams
    for each beam from the initial set
        call function trace_beam(scene, beam)
end_algorithm
```

In our designed system, the **trace_beam** function was implemented as follows:

1. All obstacles within the volume of the beam *B1* to trace are stored into a list ordered by their increasing distance from the beam's apex.
2. Each obstacle *F* from the list (from the nearest to the farthest) is tested to see whether any of its part occludes a part of the *B1* beam volume but is not yet occluded by any other obstacle. If it does, a

reflecting beam, *B2*, is created and inserted into the beam tree as a child of *B1* and is of the following properties:

- The apex of *B2* is obtained as the planar reflection of the *B1*'s apex about the plane of the obstacle *F*.
- The base shape of *B2* is obtained as the intersection of the projection of the *B1*'s base into the plane of the obstacle *F* and the obstacle *F*.

3. For each newly created beam *B2* from the children of *B1*, the **trace_beam** function is performed if a limit has not been reached (e.g., the beam *B2* is of smaller depth in the tree than a certain limit, the energy represented by *B2* is greater than a certain limit, etc.).

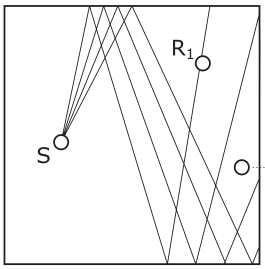


Figure 2. *Ray tracing.*

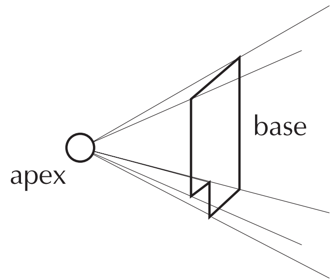


Figure 3. A beam.

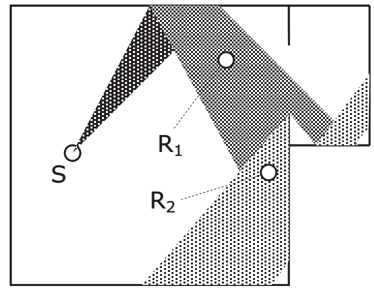


Figure 4. *A single beam; S –source, R_1, R_2 –receivers.*

2.3.3. The statistical approaches. The human auditory system is able only to distinguish the early echoes. As the late reverberation phase only gives the information about the size of the environment, it is possible to model the late reverberation phase using a statistical model where the echoes contributing to the simulated IR are randomly generated.(Funkhouser, et al., 2002).

2.3.4. The Convolution. The process of applying the IR to the sound signal for spatialization is usually modelled as the convolution of the sound signal and the IR. The convolution of two discrete signals in the Digital Signal Processing (DSP) is usually defined as:

$$f_1[t]*f_2[t]=\sum_{u=0}^{t} f_1[u]-f_2[t-u]; \qquad f_1,f_2 \text{ are the input signals} \qquad (1)$$

3. THE SPATIAL SOUND SYSTEM

The main function of the spatial audio system we designed is to perform off-line (non real-time) simulations of the sound propagation between sources and receivers, taking into account the acoustic response of the environment. Our system employs a hybrid sound propagation model built as a combination of a beam tracing algorithm (for the early echoes) and a statistical model (for the late reverberation).

The process of modelling the acoustic response of the environment consists of the two fundamental steps:

1. IR is computed as a result of simulation of the propagation of the sound from the source to the receiver in the environment.
2. The sound signal representing the acoustic activity of the sound source is taken to the convolution with the IR obtained in the previous step. The result is the spatialized audio signal.

3.1 The Architecture of the Spatial Audio System

The architecture of the designed system is described in Figure 5. The main module of the system is the Task Controller. The Task Controller initiates the calculation by reading the configuration of the rendering task from an external task description file. The format of the task files was chosen to be a proprietary derivative from the XML language for easy readability by both humans and computers. After all parameters are fetched (i.e., the scene description, the position and sound stimuli of the sound sources, the position of the sound receivers, and the acoustic properties of materials), they are stored into the Scene Representation module.

The Task Controller then activates the Sound Propagation Model module whose main purpose is to perform the simulation of the sound propagation in the environment of the scene. The module implements both the beam tracing and the statistical reverberation techniques. It is done by using the scene description (room size and texture, the obstacles' locations, etc) and the location of the sources and receivers to compute the IR profile (i.e., the individual echoes). The IR profile is stored in the Scene Representation module.

Finally, the Convolution module performs the spatialization of the raw sound signals using the echoes generated by the Sound Propagation Model module. The results are written to the output sound files through the Output Channels. An Output Channel is assigned to each elementary receiver. The resulting sound file is stored into an external output file of a standard format, which is presented through selected target device.

In the designed system, a multi-channel model of the sound receivers is employed. This model assumes that every receiver consists of several elementary receivers (e.g., a headphone consists of 2 elementary receivers: the right and the left receivers). Each elementary receiver contains a unique directivity filter. This model enables the system to render spatial audio signals for multiple target devices, e.g., headphones (where the HRTF of the listener is used as the directivity filters) or multi-speaker systems (quadraphonic, 5.1, etc.).

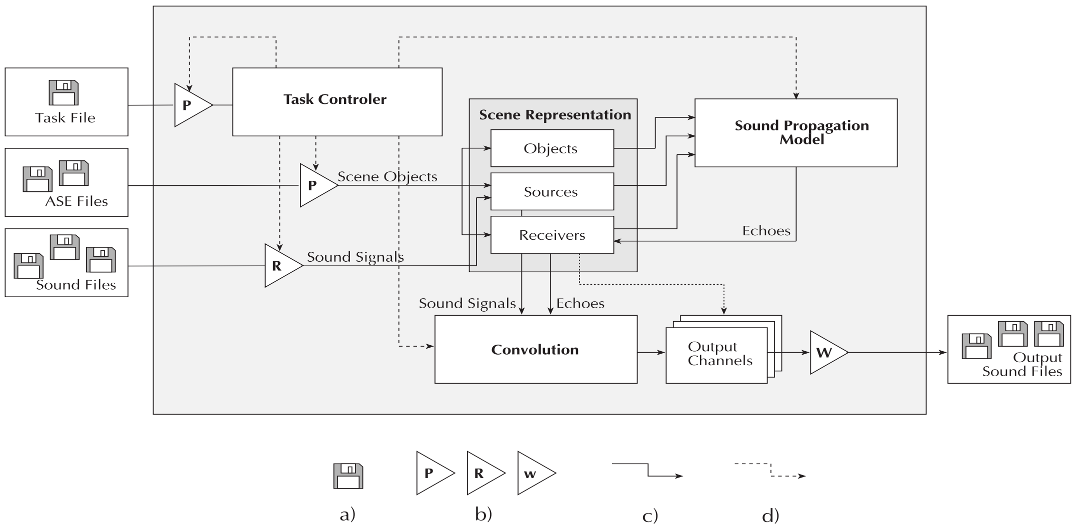


Figure 5. *The architecture of the designed system. a) Permanent data structure (file), b) Parser, reader of binary files, writer of binary files, c) Data flow in the direction of arrow, d) Spawn of a new instance of a block pointed to by the arrow.*

Because of the performance issues and the limited availability of the required 3rd party libraries, the system was implemented in the C++ language using Microsoft Visual C++ 6.0. The modules that involve DSP were implemented using the Intel Signal Processing Library version 4.5 (Intel, 2000) as it offers useful signal processing primitives such as the superposition of signals, FIR filtering, convolution, etc. The binary operations with polygons were implemented using The General Polygon Clipping Library (Murta, 2003).

4. USER EVALUATION

A group of prospective users has evaluated our spatial audio system and the fit of the algorithms used in it.

4.1 The Stimuli

We have acquired two different sets of stimuli: The sounds in the real scenes (the recorded scene stimuli) and the sounds spatialized by our system (the simulated scene stimuli.)

The real scene sounds were recorded using a stereophonic microphone PHILIPS SBC 3050 and a SoundBlaster 16 compatible sound card. Seven distinct sounds: guitar, flute, mobile phone ringing, human voice, cane tapping, glass tinkling and handclapping, were recorded in three different room conditions, coded small (S), medium (M) and large (L). The characteristics of these rooms are summarized in Table 1.1.

To create the spatialized sounds, the dry sounds (the signals without the effects of the surrounding environment) were recorded in a music studio with a very short reverberation time (less than .05s). We have used the AKG C1000S microphone and the Midiman Delta 1010 sound card. The sound signals were stored as a set of 44.1 kHz PCM files. Later on, the effects of the environments were simulated using the designed spatial audio system. It was a two-step process:

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

1. A model of the real rooms was created using the 3D Studio Max (and stored in its native format ASE.)
2. Each room model has been used with each recorded dry sound signal as an input to our system.

Table 1. *The approximate characteristics of the real scenes*

Environments	Dimensions	Surfaces	Reverberation length
Bedroom (S)	4 × 4 × 2.5 m	Plaster, carpet, wood	.2 s
Hallway (M)	8 × 3 × 5 m	Plaster, marble	1 s
Stairway (L)	12 × 12 × 10 m	Plaster, marble, tiles	3.5 s

4.2 The Evaluation Method

Nine registered blind participants (8M, 1F; mean age 29.3 with a S.D. of 6.76 years) listened to 42 sound files (7 sound types x 3 environments x 2 scenes) through a headphone. The sequence of the sounds played was controlled so that no adjacent sounds shared any similarity (e.g., the following stimulus to the flute sound in a simulated small room had to be a recorded scene, being not a flute sound, nor a small room). Each participant went through two sessions of evaluations with no other participant around. In the first session, each participant listened to one of the two sets of sounds. The order of the second set of sounds is the reversed order of the first set. After listening to each sound, the participants answered in writing three questions:

1. What sound was it?
2. Was that sound more likely to be from a small (S), medium (M) or large (L) room?
3. Was that room more likely to be a real room (R) or simulated using computer (C)?

After taking a short break, each participant listened to pairs of sounds in the second session (these pairs represent all possible combinations of room sizes and scenes, e.g., flute sounds in recorded S and M rooms, handclaps in recorded and simulated L room etc). Each participant listened to either one set of sound pairs or its reversed order. After listening to each pair of sounds, they answered in writing two questions:

1. Did you hear any difference between sound 1 and 2? Y/N
2. If Y, describe the difference.

4.3 Results and Analysis

4.3.1. The first session. The first question in the first session was intended to encourage the participants to listen carefully. Therefore, in this paper the answers were not analysed.

The answers to the second question were scored 0, 0.5 or 1. When the participants answered correctly, they were scored 1. A score of 0.5 was given when the difference between the correct and the wrong answers was one room size (e.g., a participant answered S for a sound in an M room). When the difference was two sizes (a participant answered S for L or vice versa) then a score of 0 was given.

The one-way Analysis of Variance (ANOVA) reveals that across all participants, the sum of scores for the room size question were not significantly different between the recorded and simulated scene groups, with $F(1,376) = 0.03$, $p = 0.862$. This might mean that the designed system successfully simulates various room sizes. Further analysis shows that the difference was not significant in any room size (see Figure 6).

The answers to the third question were scored 0 (wrong) or 1 (correct). In essence, this question asked the participants to perform a signal detection task. Signal detection theory (SDT) is a method of assessing the decision making process of detecting two classes of item (in this study, R and C). In an SDT task, the answers are described in terms of hit (H) and false alarm (FA) rate. In this study, the hit rate is defined as the proportion of the correct C answers in the C room conditions as explained in Table 2.

The most commonly used SDT measure is d' (discriminability), which is the standardized difference between the means of the distribution of the two classes. Larger absolute value of d' means higher distinctiveness between the two classes and d' near zero indicates chance performance. The maximum value of d' is around 5. The formula of d' is $d' = z(H) - z(FA)$ where z is the performance in terms of the number of standard deviations above or below the mean. Table 3 lists the H and FA or each room condition.

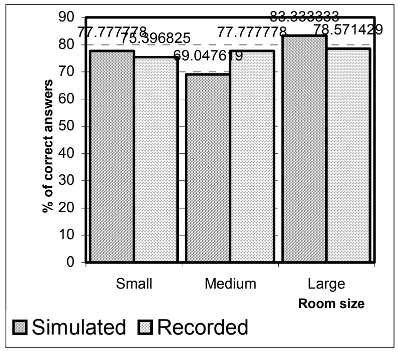


Figure 6. *The percentage of correct answers to the room size questions.*

Table 2. *The decision matrix of the user test.*

	C room	R room
Answer = C	Hit (H)	False Alarm (FA)
Answer = R	Miss (M)	Correct Rejection (CR)

Table 3. *Hit, false alarm and d' for each room condition.*

	S	M	L
H (%)	49	41	71
FA (%)	17	41	43
d'	0.935	0	0.38

Table 3 shows that in the medium room condition, the participants were not able to distinguish between the recorded and simulated scenes. The participants were most able to distinguish between those scenes in small room condition, followed by the large room condition. However, even in the latter two room conditions, the d' values were still quite small. These results might mean the designed system was successful in modelling the medium environment (hence, the simulated and recorded scenes could not be distinguished), and was less successful in simulating the small or large room conditions, although taking the absolute values of d', there was some degree of success in simulating these two conditions.

Looking at the types of sounds, there was no significant difference between various sounds in terms of the scores of the room size ($F(6,371) = 0.498$, $p = 0.106$) or the nature of scenes ($F(6,371) = 1.763$, $p = 0.810$). However, some interesting observations emerged from the sounds used to question the participants.

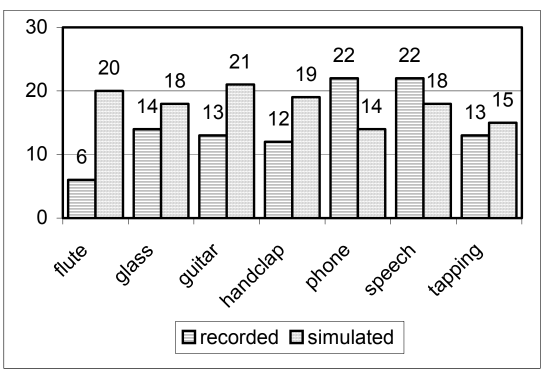


Figure 7. *The scores for the recorded vs. simulated scene question by type of sound.*

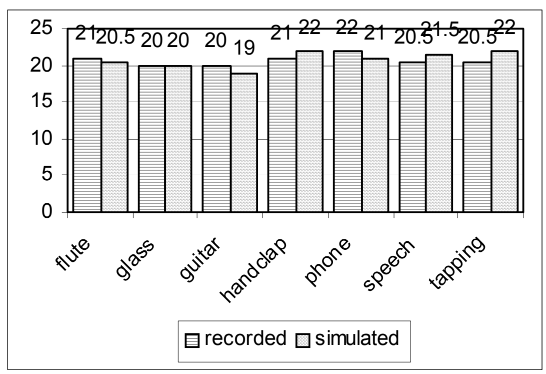


Figure 8. *The scores for the room size question by type of sound.*

Figure 7 and 8 depict the scores for each sound. Figure 10 shows that when the participants were asked whether the sounds were from recorded or simulated scenes, the highest scores were from the recorded phone ringing and human's speech sounds (both were 22 out of 27), while the highest for the simulated scenes were from the guitar and flute sounds. This finding may indicate that when musical instruments were used, the participants could detect that the sounds came from simulated scenes quite accurately. Similarly, when the sounds are ones that the blind people may hear in their daily life (in this case, the human voice and the phone ringing), they were able to detect quite accurately that these sounds came from real environments.

Figure 7 shows that the score from each sound was very similar when the participants were asked about the room size in both the recorded and simulated scenes (the scores ranges from 19-22 out of 27). This may simply mean that the types of sounds do not have any effect in helping or hindering blind people in guessing the size of room in both recorded and simulated scenes. In other words, the type of sounds does not affect users' perception of room sizes in both room scenes.

4.3.2. The second session. The second session aims to gather ideas about users' terms when describing the differences between various experimental conditions. In general, most answers to the questions about the differences between the sounds in different room sizes contain the words: echoes, reverberation, resonance, closer/further and depth. Most answers to the question about the difference between the sounds in simulated vs. recorded scenes contain the words: more/less natural/realistic, crispier/sharper and metallic/ digital sounds. As these words represent the terms that blind users may use to describe the sound variations in different environments, these terms may be useful in the context of the O&M training for blind people (e.g., "Please listen carefully to the sound variation. The first sound is from a larger room than that of the second one as the first sound has more resonance.")

5. CONCLUSIONS

The results of the user studies indicated that the algorithms behind the designed spatial audio system were able to simulate the environments to certain extent. The system was able to simulate the sound variation in different room sizes successfully, indicated by the lack of significant differences between the sum of scores in the simulated and recorded scene groups. When simulating a medium room, the difference between the reverberation of the simulated and recorded scenes were not noticeable to our blind participants. Based on these results, we can speculate that the designed audio system is potentially useful as a part of the O&M training suite for blind and visually impaired people, preferably to simulate sounds in a medium room. Our immediate future work is to implement a scene audibility graph, that would describe the acoustic relations among different places in the scene, regardless on the actually employed method of the spatial audio simulation. We also plan to integrate this system into a training suite and testing the suite with its prospective users. Further studies are also needed to investigate users' mental model of various room conditions.

6. REFERENCES

V R Algazi, R O Dudaa, R Duraiswami, N A Gumerov and Z Tang (2002), Approximating the head-related transfer function using simple geometric models of the head and torso, *Journal of Acoustics Society of America* **112**, pp. 2053-2064.

C Colwell, H Petrie, D Kornbrot, A Hardwick and S Furner (1998), Haptic Virtual Reality for Blind Computer Users, *Proc. ASSETS 1998*, pp. 92-93.

R O Duda, C Avendado and V R Algazi (1999), An Adaptable Ellipsoidal Head Model for the Interaural Time Difference, *Proc. ICASSP 1999*, pp. 965-968.

K M Franklin and J C Roberts (2003), Pie Chart Sonification, *Proc. of 7th International Conference on Information Visualization*, pp. 4–9.

T Funkhouser, J M Jot and N Tsingos (2002), Sounds Good to Me!, Computational Sound for Graphics, Virtual Reality, and Interactive Systems, *SIGGRAPH 2002 Course Notes*. http://www.cs.princeton.edu/gfx/papers/funk02course.pdf.

T Funkhouser, P Min, I Carlbom (1999), Real-Time Acoustic Modeling for Distributed Virtual Environments, *Proc. SIGGRAPH 1999*, pp. 365–374.

P S Heckbert and P Hanrahan (1984), Beam Tracing Polygonal Objects, *Computer Graphics* **18**, *3*, pp. 119–127.

D P Inman and K Loge (1999), Teaching orientation and mobility skills to blind children using simulated acoustical environments, *HCI* **2**, pp. 1090-1094.

Intel® Signal Processing Library (2000), *Software library documentation 630508-012*.

H Kuttruff (2000), In Room Acoustics, 4th ed., Spon Press, London, U.K.

T Lokki, L Savioja, R Vaananen, J Huopaniemi and T Takala (2002), Creating Interactive Virtual Auditory Environments, *IEEE Computer Graphics & Applications* **22**, pp. 49-57

A Murta (2003), A General Polygon Clipping Library. http://www.cs.man.ac.uk/aig/staff/alan/software/gpc.html

Y Seki and K Ito (2003), Study on Acoustical Training System of Obstacle Perception for the Blind. In *Assistive Technology - Shaping the Future* (Craddock, McCormack, Rielly & Knops, Eds.), pp. 461-465.

The Royal Blind School (2003), Orientation and Mobility. http://www.royalblindschool.org.uk/Departments/Mobility.htm

AudioMath: blind children learning mathematics through audio

J H Sánchez and H E Flores

[1]Department of Computer Science, University of Chile,
Blanco Encalada 2120, Santiago, CHILE

jsanchez@dcc.uchile.cl, *hflores@dcc.uchile.cl*

ABSTRACT

Diverse studies using computer applications have been implemented to improve the learning of children with visual disabilities. A growing line of research uses audio-based interactive interfaces to enhance learning and cognition in these children. The development of short-term memory and mathematics learning through virtual environments has not been emphasized in these studies. This work presents the design, development, and usability of AudioMath, an interactive virtual environment based on audio to develop and use short-term memory, and to assist mathematics learning of children with visual disabilities. AudioMath was developed by and for blind children. They participated in the design and usability tested the software during and after implementation. Our results evidenced that sound can be a powerful interface to develop and enhance memory and mathematics learning in blind children.

1. INTRODUCTION

Audio-based interfaces to foster cognition in blind children have been increasingly developed for people with disabilities. A number of studies have focused on the design of audio-based interfaces and evaluated their impact on learning and cognition (Baldis, 2001; Lumbreras, & Sánchez, 1998; McCrindle & Symons, 2000; Mereu & Kazman, 1996; Sánchez, 2000; Sánchez, 2001; Tan, 2000, Winberg & Helltrom, 2000).

Most of these studies are based on interactive software that cannot be fully adapted to their needs and requirements. Actually, most of them are fixed prototypes without enough flexibility and versatility.

A seminal study developed the first proof-of-concept application to navigate virtual environments through audio by blind children (Lumbreras & Sánchez, 1998). The study used audio to enhance spatial cognitive structures and found that spatialized sound can be used to develop spatial navigation skills in a virtual environment.

Some studies designed experiences with audio stimuli to simulate visual cues for blind learners (Mereu & Kazman, 1996). As a result, they found that by using 3D audio interfaces blind people can help to localize a specific point in a 3D space. They performed with precision but slower than sighted people concluding that navigating virtual environments with only sound can be more precise to blind people in comparison to sighted persons (Mereu & Kazman, 1996).

Other studies describe the positive effect of 3D audio-based virtual environments (Cooper & Taylor, 1998). A study in the same line of research used sensory virtual environments through force feedback joysticks simulating real places such as the school or work place. They probed the hypothesis that providing appropriate spatial information through compensatory channels can improve the performance of blind people (Lahav & Mioduser, 2000).

A research work in the same track of (Lumbreras & Sánchez, 1998) concluded that a traditional computer game such as Space Invader can be replicated with 3D sound. Researchers used force feedback joysticks as input interface by letting to play blind to sighted children to share the same experience (McCrindle & Symons, 2000). An interesting study tested blind and sighted people with covered eyes across audio stimuli by tracing specific places through sound. The skill to hold in mind the specific localization without concurrent perceptual information or spatial update was evaluated (Loomis et al., 2002).

Studies focusing the development of computer applications to develop and enhance memory through virtual environments in people with disabilities are scarce. Some memory processes such as spatial distribution through virtual reality environments were developed by active participation of people with disabilities (Attree et al., 1996). As a result, active participation enhanced memory for spatial layout whereas passive observation enhanced object memory.

Other studies have also used immersive virtual reality to analyze spatial memory in patients with brain damage. They developed a test as proof-of-concept of the critical role that can be played by virtual reality as experimental tool for memory purposes (Morris et al., 2000). The use of virtual reality was also found to be relevant to improve the validity and reliability of neurological evaluation and rehabilitation (Rizzo et al., 2002). The impact of spatial audio on memory was analyzed concluding that audio stimuli improves memory and perceived comprehension by choosing spatial audio instead of non spatial (Baldis, 2001).

A trend in the literature is the absence of long-term usability studies. Increasing voices in the field are sustaining that spatial audio may have a reduced impact when it is not associated to specific cognitive tasks. Besides to the need for sound-based virtual environment studies there is a demand for more rigorous and systematic usability studies by and for children with visual disabilities.

A most recent study have developed a sound-based virtual environment for children with visual disabilities to enhance memory and fully tested the software with diverse children concluding that it can help to enhance the development and practice of short-term memory in children with visual disabilities. (Sánchez and Flores, 2003, 2004).

Learning the basis of mathematics has been a current issue in literacy literature. Most studies worldwide agree that children do not learn mathematics adequately in the early grades. This has a tremendous impact on further learning. In a world heavily based on science and technology children without understanding the basics of math limit their role in the society. Children with visual disabilities are not the exception. Actually in many respects this issue is radicalized in these children. When blindness is associated to social deprivation the issue of learning primary school mathematics is really a critical issue (Edwards & Stevens, 1993; Sahyun et al., 1998; Scadden, 1996). Thus one of the greatest challenges for children with visual disabilities has been the learning of, and access to, mathematics information (Sahyun et al., 1998). Early learning and practice may project a better construction of mathematics knowledge in visually impaired children (Edwards & Stevens, 1993).

In this study we intended to foster learning and practice of mathematical concepts such as positional value, sequences, additive decomposition, multiplication, and division. Through opening pairs of tokens in a board with several levels of difficulty, the child has to find the corresponding pair of tokens in accordance with the mathematical content. Thus AudioMath is used to assist learners with visual disabilities in learning concepts such as establishing correspondence and equivalency relationships, memory development, and differentiating tempo-spatial notions.

2. DESIGN OF AUDIOMATH

2.1 Model

AudioMath has different components. *Specific content* models the representation problem to generate a grid with a pair of related tokens between them to be solved by the child. *Random card generator* is the editor that allows setting the level of complexity and contents from a gallery. *Computer model* is the representation of the real problem. It includes state system variables such as number of correct token pairs, time, and score as well as parametric variables such as level, content, and user name. *Projection* implies transforming input signals to changes perceived by blind users either audible or tactile. It bridges system and interfaces through bidirectional feedback from and toward the user actions. *Interface* includes input/output interaction devices such as audio, keyboard, mouse, force feedback joystick, and tablets.

The model is based on a matrix with rows and columns. There are four levels of complexity: Level 1 with four tokens (two rows and two columns), level 2 with six tokens (three rows and two columns), level 3 with twelve tokens (three rows and four columns), and level 4 with sixteen tokens (four rows and four columns). Colours are used for children with residual vision (Rigden, 1999). This model meets the minimum standards proposed by (Sánchez et al., 2004) for software design, and thus validating AudioMath as an appropriate virtual environment to be used for the learning of children with visual disabilities.

2.2 Software and hardware Tools

AudioMath was developed by using Macromedia Director 8.5 and a library of routines for external joystick control, Xtra RavJoystick. A joystick Side-Winder was used in conjunction with Force Feedback, mouse, keyboard, and Wacon tablets.

2.3 *Interfaces*

The principal interfaces of AudioMath for children with residual vision are displayed in Figure 1. For blind children interfaces are only audio-based. A is the identification screen with two modes: facilitator or student (two buttons). B considers the level of complexity (list box), content (list box), and input device (buttons). Content can be filled, upgraded, and edited by using different media. C is the main interface of AudioMath and includes options such as the position of the card grid, accumulated score, time elapsed (through speech), restart, register, and exit (through buttons). D includes a logging actual use register (buttons) describing each game and movements.

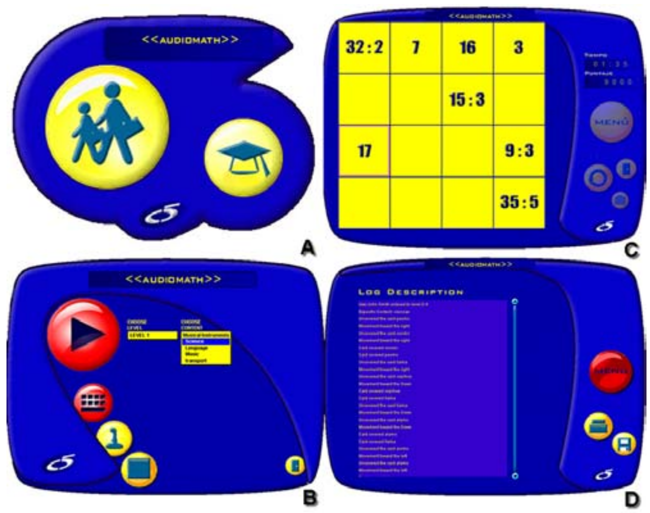


Figure 1. *Interfaces of AudioMath*

2.3 *Interaction*

AudioMath allows interaction with available interface elements such as buttons and text screens through keyboard. Each interaction triggers an audio feedback and a high visual contrast screen to be perceived by children with residual vision. The child has to move through a grid and open the corresponding token. Each cell has associated music tones that identify the position in the grid. Sound is listened when moving through cells. When open a token the associated element is visible by triggering an audio feedback. For example, if the image is a car, a real traffic car sound is triggered. When open a correct token pair a feedback is set. Finally, when all pairs are made the time used, the final score, and feedback are given.

AudioMath can be interacted through keyboard, joystick, and tablets. A few keystrokes are used with the keyboard. These devices have been used in different applications developed by the research group after testing with children with visual disabilities. The Microsoft SideWinder joysticks in conjunction with Xtra RavJoystick for Macromedia Director allow grading the user position in the grid and give direct feedback with different forces. Counter forces to the movement are generated per each token position change as well as vibratory forces indicate to be near to the grid edge: up, down, left, and right. Force Feedback Joysticks allow direct interaction with diverse degrees of freedom. A plastic graphic grid is posed on the tablet defining the position of each token. A pen is used to point and select interface elements.

3. COGNITIVE EMPHASIS

AudioMath was designed to enhance memory in blind children and children with residual vision. For reading and writing skills learners have to have developed visual and audio memory to discriminate graphemes and phonemes. This memory prevents to be confused by graphemes for spatial orientation as well as by phonemes with similar sounds. AudioMath allows children with visual disabilities to practice and rehearse their short-term memory.

These children practice audio memory (blind and children with residual vision) and visual memory (children with residual vision). The tasks implied to exercise audio/oral, visual/oral, audio/graphic, and visual/graphic memory. The software also emphasizes learning concepts such as establishing correspondence and equivalency relationships, memory development, and differentiating tempo-spatial notions.

AudioMath was implemented to go further than just enhancing general domain skills such as memory and tempo-spatial notions by integrating mathematics content based on the current school curriculum. We embedded the software with mathematical concepts such as positional value, sequences, additive decomposition, multiplication, and division. We wanted to observe how audio-based virtual environments can foster the construction of mathematics learning in the mind of children with visual disabilities.

4. METHODOLOGY

4.1 Participants

The study was developed with ten children ages 8 to 15 who attend a school for blind children in Santiago, Chile. The sample was conformed of 5 girls and 5 boys. Most of them have also added deficits such as diverse intellectual development: normal, slow normal, border line, below to normal, and with mental deficit. Four special education teachers also participated. All learners met the following prerequisites: to be legally blind, to know the natural numbers, to express sequences orally, to order numbers, to decompose numbers through audio means, to mentally estimate results of additions and subtractions, to mentally determine products and coefficients, to mentally decompose numbers in additions, to manipulate multiplication tables efficiently, and to have notions of fractions.

4.2 Evaluation Instruments

Three measurement tests were used to evaluate the impact of AudioMath on learning and practice of mathematical concepts such as positional value, sequences, additive decomposition, multiplication, and division. Immediate audio memory test (Cordero, 1977), evaluation of mathematics knowledge test (Chadwick & Fuentes, 1980), and a usability evaluation test for end-users. The immediate audio memory test has the purpose to measure logic, numeric memory, and associative memory from audio stimuli. The evaluation of mathematics knowledge test measures: 1. The capacity to understand numbers (oral and written); 2. The skills to make oral and written calculations; 3. The skills to count numeric series and graphic elements; and 4. The skills for mathematic reasoning.

4.3 Procedure

Children were tested in the school from July to November 2003, twice a week, two one hour sessions per week. They followed the steps of pre-testing (immediate audio memory and evaluation of mathematics knowledge tests), interacting with AudioMath, solving cognitive tasks (see Figure 2 and 3), and post-testing (immediate audio memory and evaluation of mathematics knowledge tests). Interacting with AudioMath and solving cognitive tasks were the main steps of the study. During these steps children were observed and assisted by four special education teachers filling check lists and registering observed behaviours. They also applied a usability evaluation test for end-users developed by the authors.

Figure 2-3. *Children solving cognitive tasks with AudioMath.*

5. COGNITIVE IMPACT

During the interactive sessions we realized that mathematical content used was appropriate to the educational level of the sample. We analyzed the results case by case because the sample was not homogeneous in key variables such as cognitive deficits and different degrees of blindness.

Children performed increasingly well in both tests: audio memory and mathematics knowledge. An overall view of initial results shows pre-test – post-test gains in mathematics knowledge (see Figure 4), thus indicating that interaction with AudioMath associated with cognitive tasks can improve mathematics learning in these children. Audio memory pre-tests – post-tests show some gains after interacting with AudioMath, thus indicating that short-term memory can be enhanced by using this software. Gains were higher in mathematics knowledge than audio memory (see Figure 5).

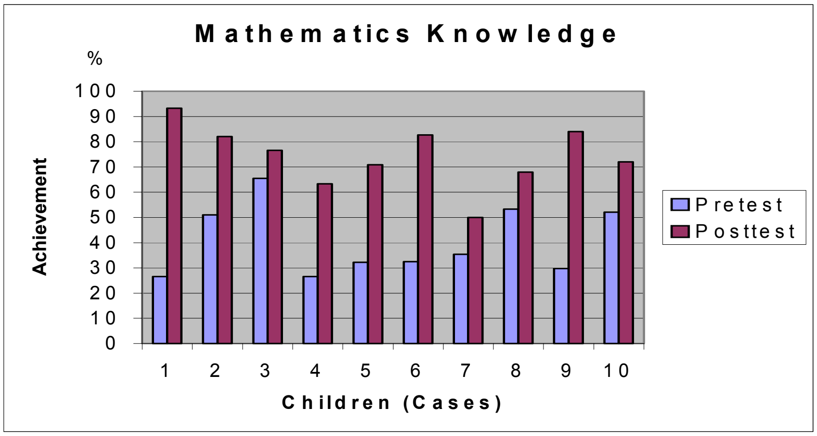


Figure 4. *Pre-test – post-test gains in mathematics knowledge*

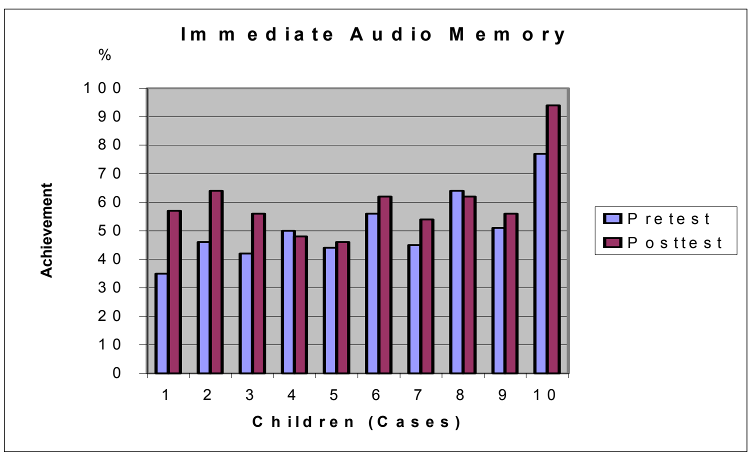


Figure 5. *Pre-test – post-test gains in short-term memory*

In mathematics learning the results were promising. Most evaluated mathematics content was well attained by learners with visual disabilities. The highest gains were in oral calculation (75%) and countdown of numeric series (100%). We believe that these results were also partially due to a better attitude of the children toward work and mathematics knowledge construction such as multiplication tables and use of mathematics operations. They impacted positively on learners by increasing their certainty.

Children also obtained high gains in short-term memory. They improved their numeric memory (69th percentile score), associative memory (65th percentile score), and logic memory (3rd percentile score). This means that they performed better than, or received a higher score than, 69%, 65%, and 3% of other learners that took the same tests.

Finally, we can conclude that thanks to the interaction with AudioMath and the associated cognitive tasks learners with visual abilities developed mathematics skills and short-term memory. This is a major result in our research because we are initially observing that audio-based virtual environments can foster the construction of mathematics learning in the mind of children with visual disabilities.

6. DISCUSSION

We have introduced AudioMath, a virtual environment to enhance memory and mathematics skills in children with visual disabilities through audio. The software was made by and for children with visual disabilities. They participated actively in the development of the software. We have also designed interfaces for both blind and children with residual vision. A usability study was implemented with end-users, facilitators, observers, and experts.

AudioMath was highly accepted by end-users. They liked, enjoyed, and were motivated when interacted with the software. The flexibility of this application is also a plus. Teachers, children, and parents can include new objects and sounds to adapt them to their needs. Thus, children with visual disabilities can choose sounds to be interacted with and embed them into AudioMath. Content can be changed and updated easily. AudioMath can be used to support memory when learning specific concepts and processes in a given subject matter. AudioMath can be used for learning primary school mathematics.

The use of concrete materials was also a plus in this study. The children's understanding was easier when they first interacted with concrete materials and then with AudioMath. Parallel interaction with both concrete material and AudioMath was also an advantage. Once they developed their own mental model of the software the interaction with AudioMath was enriched.

Force Feedback Joysticks introduced a new scenario in virtual environment for blind children. They can provide information and tactile sensations through force feedback. This can help to decrease audio stimuli and relief possible acoustic pollution. Joysticks are devices with a high potential of use due to the availability of many buttons.

Our model fits well the learning of primary school mathematics concepts such as positional value, sequences, additive decomposition, multiplication, and division. Children performed increasingly well in both tests: audio memory and mathematics knowledge. Oral calculation and countdown numeric series were highly achieved as well as numeric and associative memory. Concrete cognitive tasks were crucial in this achievement. We firmly conclude that interaction with AudioMath associated with cognitive tasks can help to improve mathematic learning in these children.

More qualitative data are being analyzed. Most of them are case study because each child with visual disabilities is a whole case that deserves a deep analysis to construct meaning about the role that can play audio-based devices in learning general and specific domain skills.

Finally, we are convinced that further research studies we are implementing right now concerning mathematic learning will reaffirm our hypothesis that audio-based virtual environment can foster the construction of mathematics learning in the mind of children with visual disabilities.

Acknowledgements: This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1030158.

7. REFERENCES

A Cordero (1977), *Immediate Audio Memory*. 5th edition. TEA Editorial, Applied psychology publications, Madrid, Spain, 1977.

A D N Edwards and R D Stevens (1993), Mathematical representations: Graphs, curves and formulas. In D. Burger and J. C. Sperandio (ed.) *Non-Visual Human-Computer Interactions: Prospects for the visually handicapped*. Paris: John Libbey Eurotext, (1993), pp. 181-194.

A Rizzo, T Bowerly, J Buckwalter, M Schultheis, R Matheis, C Shahabi, U Neumann, L Kim and M Sharifzadeh (2002), Virtual environments for the assessment of attention and memory processes: the virtual classroom and office. In *Proceedings of the 4th International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT 2002*, 18-20 September 2002, Veszprém, Hungary.

C Rigden (1999), The eye of the beholder-designing for colour blind. *British Telecommunications Engineering*, Vol. 17, January.

C Sjostrom (2001), Using haptics in computer interfaces for blind people. *Proceeding of the ACM CHI ΄01*, Seattle, Washington, USA, March 31 – April 5, 2001, Vol 3, 1. pp. 245-246.

E Attree, B Brooks, F Rose, T Andrews, A Leadbetter and B Clifford (1996), Memory processes and virtual environments: I can't remember what was there, but I can remember how I got there. Implications for people with disabilities. In *Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies, ECDVRAT* 1996, 8-10 July 1996, Maidenhead, Berkshire, UK.

F Winberg and S Helltrom (2000), The quest for auditory manipulation: the sonified Towers of Hanoi. In *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies*, *ICDVRAT 2000*, Sardinia Italia, 23-25 September, 2000. pp. 75-81.

H Tan (2000), Haptic interfaces. *Communications of the ACM, 43(3)*. pp. 40-41.

J Baldis (2001), Effects of spatial audio on memory, comprehension, and preference during desktop conferences. In *Proceeding of the ACM CHI 2001*, Seattle, Washington, USA, March 31 – April 5, 2001. Vol 3, 1. pp. 166-173.

J Loomis, Y Lippa, R Klatzky amd R Golledge (2002), Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(2),* pp. 335-345. 2002

J Sánchez (2000), 3D interactive games for blind children. In *Proceedings of Technology and Persons with Disabilities, CSUN 2000*. Los Angeles, USA.

J Sánchez (2000), Usability and cognitive impact of the interaction with 3D virtual interactive acoustic environments by blind children. In *Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies*, *ICDVRAT 2000*, Sardinia Italia, 23-25 September, 2000. pp. 67-73.

J Sánchez (2001), Interactive virtual acoustic environments for blind children. In *Proceedings of ACM CHI ´2001*, pp. 23-25. Seattle, Washington, USA, March 31 – April 5, 2001.

J Sánchez and H Flores (2003), AudioMemory: Developing memory in children with visual disabilities through audio. *Proceedings of the 8th International Workshop on Educational Software, TISE 2003*, Santiago, Chile, November, 2003.

J Sánchez and H Flores (2004), Memory enhancement through audio. To be published by in the *Proceedings of ACM Assets 2004*, Atantla, Giorgia, USA, October 18-20.

J Sánchez, N Baloian and H Flores (2004), A methodology for developing audio-based interactive environments for learners with visual disabilities, *Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunications*, *ED-MEDIA 2004, Lugano, Switzerland, June 21-26.*

L Scadden (1996), Making mathematics and science accessible to blind students through technology. *Proceeding of Resna 96 Annual Conference*, (1996), June 7-12, Salt Lake City, Utah, pp. 51-58.

M Chadwick and M Fuentes (1980), *Evaluation of Mathematics Knowledge* (adaptation of Benton-Luria test). Santiago, 1980.

M Cooper and M E Taylor (1998), Ambisonic sound in virtual environments and applications for the blind people. In *Proceedings of the Second European Conference on Disability, Virtual Reality, and Associated Technologies, ECDVRAT 1998*, Skövde, Sweden, 10-11 September, 1998. pp. 113-118.

M Lumbreras and J Sánchez (1998), 3D aural interactive hyperstories for blind children. *International Journal of Virtual Reality 4(1. pp.,* 20-28.

O Lahav and D Mioduser (2000), Multisensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping, orientation, and mobility skills. In *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT 2000*, Sardinia Italia, 23-25 September, 2000. pp. 53-58.

R McCrindle and D Symons (2000), Audio space invaders. In *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies*, *ICDVRAT 2000*, Sardinia Italia, 23-25 September, 2000. pp. 59-65.

R Morris, D Parslow and M Recce (2000), Using immersive virtual reality to test allocentric spatial memory impairment following unilateral temporal lobectomy. In *Proceedings of he 3th International Conference on Disability, Virtual Reality and Associated Technologies, ICDVRAT 2000*, Sardinia Italia, 23-25 September, 2000.

S Mereu and R Kazman (1996), Audio enhanced 3D interfaces for visually impaired users. *Proceedings of CHI'96*, ACM Press. (1996).

S Sahyun, S Gardner and C Gardner (1998), Audio and Haptic Access to Math and Science - Audio graphs, Triangle, the MathPlus Toolbox, and the Tiger printer. *Proceedings of the 15th IFIP World Computer Congress*, Vienna, September 1998, pp. 78-86.

Creating aesthetically resonant environments for the handicapped, elderly and rehabilitation: Sweden

A Lewis-Brooks[1] and S Hasselblad[2]

[1]Computer Science Department, Aalborg University Esbjerg,
Niels Bohrs vej 8, 6700 Esbjerg, DENMARK

[2]Special Needs Teacher, Emaljskolan,
Emaljgatan 1C, S-26143 Landskrona, SWEDEN

tonybrooks@cs.aue.auc.dk, stefan.hasselblad@mila.landskrona.se

[1]*www.aue.auc.dk,* [2]*www.emaljskolan.se, www.bris.ac.uk/carehere.htm*

ABSTRACT

This contribution expounds on our prior research, where interactive audiovisual content was shown to support Aesthetic Resonant Environments with (1) brain damaged children -extended here with the addition of (2) learning disabled, Parkinson's disease, and the aged. This paper appraises the experiments involved in preparing, developing and authenticating 'aesthetic resonance' within the Swedish partners' research (1 & 2). It reports on the inductive strategies leading to the development of the open architectural algorithms for motion detection, creative interaction and analysis, including the proactive libraries of interactive therapeutic exercise batteries based on multimedia manipulation in real-time.

1. INTRODUCTION

European Project IST-20001-32729, CAREHERE; *Creating Aesthetically Resonant Environments for the Handicapped, Elderly and Rehabilitation* was funded under Framework V IST Key Action 1 supporting the programme for 'Applications Relating to Persons with Special Needs Including the Disabled and Elderly.' It implicated a consortium of seven members from four countries. Following the project overview, this paper informs specifically on the investigation of the partners from Sweden who had responsibility for the 'user group programmes research' and the 'user forum network,' P3 and P7 respectively.

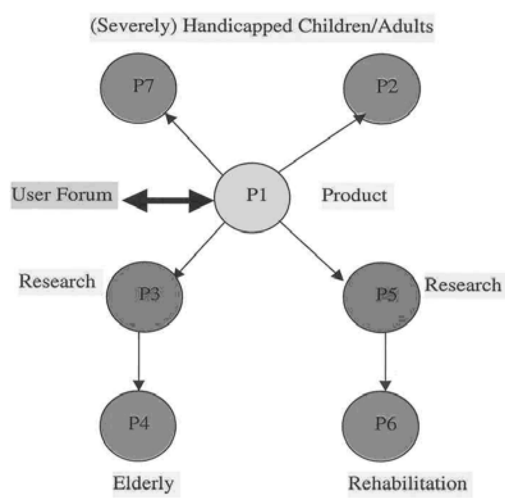


Figure 1. *The CAREHERE consortium with Swedish members P3 Lund & P7 Landskrona.*

 ISBN 07 049 11 44 2

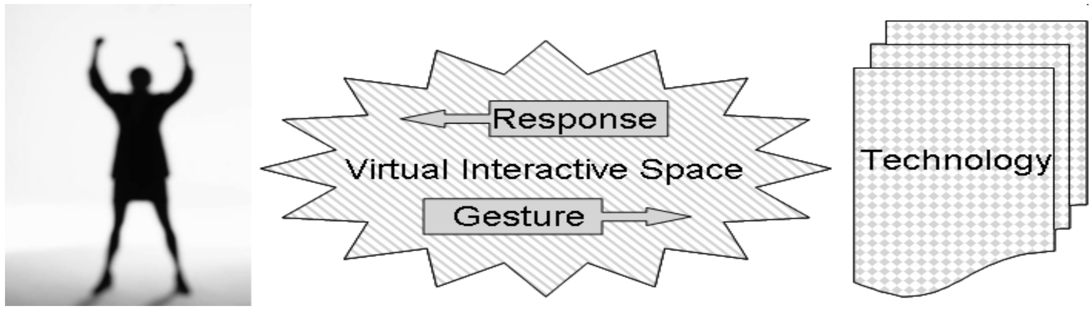


Figure 2. *Aesthetic Resonance refers to a situation when the response to intent is so immediate and aesthetically pleasing as to make one forget the physical movement (and often effort) involved in the conveying of the intention. See also – VIS* (Brooks 1999).

2. AESTHETIC RESONANCE, THE CONCEPT & THE COMMUNITY

Augmented real-world environments with inherent responsive technology are created so as to explore 'Aesthetic Resonance.' These are explored proactively where a sensory stimulating *feedback* (e.g. sound or visual), responding to an initial physiological gesticulation, instigates suggested potential closure of the afferent-efferent neural loop through subsequent *feed-forward* (also known as homeostatic control system) iteration which is anticipating the result of the associated motor action on the *feedback*. This stimulating *feedback* is selectable depending on user preferences, desires, and abilities; and similarly any *feed-forward* limitations – physiological or psychological - of the user need to be accountably optimised through the adaptability of the system so as to motivate the iterative causal interaction. This evokes a clear (at suggested various levels of consciousness) understanding of the causality involved and is analogous to the *flow state* involved in play & game psychology - exhibited when a child is engrossed in a computer game. However in our research, whilst using a computer workstation for signal processing, there is no joystick, mouse or tablets involved, nor are there attachments worn by the user such as in motion analysis laboratories that cater for rehabilitation therapy. A strategy of no invasive techniques or strict homogeny such as in most motion tracking techniques is maintained, the user is free to move in 3D Virtual Interactive Space (VIS), unencumbered by wires or attachments (Brooks 1999). This freedom is a catalyst of the concept - especially in respect of when the effort in the achieving of a goal means the overcoming of a pain.

2.1 The lineage to the scientific community's current growing interest.

The concept of using interactive feedback with movement has been around for some time. In 1990 a Danish national newspaper reported on the pioneering work of a Canadian (Rokeby 1990) where a camera was used to capture movement to play music on a computer. Later he worked within the disabled community with his system. Two years later Jaron Lanier, in his keynote address for the 'Virtual Reality and Persons with Disabilities' conference in California State University emphasized the importance for the VR industry focus to be on the *human-centred* and *empowering of the human* attributes of the technology and further informed of the incredible overlap between the communities stating, "In the history of Virtual Reality development, the community of researchers building the Virtual Reality machines and Virtual Reality software has been, in many cases, almost the same community as the people working on tools for disabilities" (Lanier 1992). In that same year Dr. Jeffery Pressing, who was a musician and an academic psychologist, published on real-time musical synthesizer techniques, improvisation composition and also numerous papers in human science. His writings [1] are a good example of the continued nascent overlap between the arts (esp. music) and technology into the human sciences. This is further exemplified in Pacchetti (*et al*. 1998, 2000) where *active* music therapy was used in rehabilitation of Parkinson's disease patients. Subsequently, publicly accessible Human Computer Interaction demonstrations and workshops (e.g. Brooks & Hasselblad 2001) presented technologies that enabled such a cross over between the communities and took advantage of the increased power and usability of computers and the inter-disciplinary research potential. All point to the fact that the lineage had come of age with the inferred growing interest from the scientific community in exploring therapeutic exercises based on interactive multimedia techniques. Further reinforcement of the interest is exemplified by the cross-disciplinary work of:- Bianchi & Saba (1998) in the autonomy of the disabled and elderly; (McComas et al. 1998) with children with disabilities; Wong & Tam (2001) in therapist-instructed training with autistic children; Brown (*et al.* 2002) in respect of social inclusion guideline for interactive multimedia; (Rutten *et al*. 2003) in supporting learning and enhancing social skills in people with autism;

(Yano *et al*. 2003, Kizony *et al.* 2003, Brooks 2004a) in respect of stroke. Continuous support of such collaborative actions [2] is essential.

3. PROJECT OVERVIEW AND BACKGROUND

This project structure was influenced by previous experiences in the ESE Caress project [2]; the FET Twi-aysi probe (Brooks *et al.* 2002) and the SoundScapes body of research (Brooks 2004b). The key areas defined for concentration in CAREHERE being (1) iterative development, (2) sonic and visual content, (3) wider user base, (4) hands off sensors, (5) quantitative evaluation, (6) user forum networking, (7) a concentration upon software not hardware, and (8) commercial exploitation. These key areas being reflected in the makeup of the consortium – with four user partners, two technical R & D partners, and one commercial partner, supported by an evolving 'User Forum' network of interested parties outside of the consortium. The earlier research investigated the use of responsive sound and visual environments as a potential supplement to therapy for the development of physical and cognitive skills through direct and immediate feedback through the aural and visual senses. The next step was to:

- Create Aesthetic Resonant Environment and Support Programme for Special Needs: supporting the development of physical and cognitive skills by interaction with a responsive audio-visual environment.
- Audio-Visual Interactive Content: Sonic Navigation, Visual Empathy, Sonic Tactility, Capture of Expressivity.
- Individual Adaptability: A unique system able to adapt to individual needs and wishes.
- Modelling and Capture of Expressivity: production of the above audio and visual content will entail innovation in both (physical) modelling and the capture of expressivity.
- Immediate Commercialisation: technical innovations being integrated into the product development of our commercial partner.

4. TECHNOLOGIES, TEETHING TROUBLES, AND METHODOLOGY

Eventual delayed testing with the infrared sensor based system developed by our coordinating & commercial partner *Personics* from Denmark gave immediate value to the iterative development process strategy as the user groups, especially our elderly users, expressed a strong preference for real/natural sounds and pictures and videos (rather than the limited synthetic sounds and abstract pictures only available with *Personics*.) There were also problems with positioning of the sensors for optimal capture and the information to feedback was not intuitive as often a movement would generate opposed and false data often crashing the system (which can be expected for a prototype system; however this was not rectified within a promised liberal timeframe) and disrupted the research. All consortium partners' realizing the failure in time ceased use of the *Personics* system in favour of *Eyesweb*, which is based on camera acquisition, and was capable of accommodating user preferences and was stable in use. As a result, *Personics* were offered the opportunity to program their sensor as an input device to *Eyesweb* so as to be usable alongside the cameras as an input sensor in our research. This they declined - a decision which was mutually acceptable as certain other features of the unstable *Personics* system were inappropriate for our work. The project thereafter was camera based, which became optimal as the cameras were used both as 'sensor' cameras for interactivity through movement with the audio-visual content and as 'observation' cameras for evaluation.

An algorithm was created by the Eyesweb software partner representatives in collaboration with the authors when they came to Lund, Sweden to see the work. This was the success story of the project as the algorithm enabled movement by a user to paint relative to a programmable (and easily changeable) velocity threshold which could adapt for user function. It used a technique called Silhouette Motion Images (SMI) which provided values that were proportional to the absolute velocity of the body part being moved. Automatic segmenting of the movement in order to distinguish between motion and pause phases was achievable with the movement cues able to be measured in real-time and associated to each motion and pause. Such data was useable to produce a symbolical description of the movements being performed. Example cues were:

- Contraction index (openness of body posture with respect to centre of gravity).
- Directness index (direct or flexible nature of a body limb trajectory).
- Durations of pause and motion phases (inspired by psychological studies).

Layered Analysis [LA] (Ellis 1996) from the earlier Caress project [2] was used adjacent to Multiple Camera Analysis [MCA] (Brooks 2002) to archive and analyze all video footage from the sessions.

5. USER GROUPS

'Kulturcentrum' Lund group (research P3) consisted of 15 cognitive impaired – learning disabled, well functioning adults with ages ranging from 20 – 42 years, able to communicate and mobile. Experimental exercises were set up where a database of abilities was archived. This ranged from working with gesticulated painting and music making (figs. 3 & 4), *Shadow-ation* (my word inc. shadows, imagination and motivation), analogue camera/monitor feedback loops (fig. 5), and dimensional environment exercises. The *Cross Modal Painting* was a follow on from earlier adjacent experiments with movement triggering image and sound; [fig. 5 in Brooks 2004d] and was positively responded to. The *Shadow-ation* was where we had the users come into the room one at a time and stand approximately 1 meter facing a white wall within a strong 2000 lumen LCD projector beam so that their shadow was strong and clear in front of them. They were instructed to create with their shadow three living items and interpret with their shadow. Over the weeks the items changed but to give an indication they would be asked for a bird, a fish and a tree. Then they were timed and videoed in response. The analogue feedback loops were only tested with the single case study below. The environment exercises were particularly fruitful as in the Centre there was an open top elevator (approx. 1.7m x 2m) between the ground floor and upper floor. The LCD projector was fixed so as to beam images down the escalator shaft and the system set up to track movement in the space. Many stated that it was like having their own disco that they could control with result of a very definite aesthetic resonance being achieved in the environment. Subsequent tests were in a large room and an auditorium with positive results.

5.1 Single case study

The single case study was set at 30 sessions each approximately 45 minutes long with a woman of 34 years. Various feedbacks were experimented in these sessions so as to ascertain what would be achievable through the system and to give subsequent feedback to the development teams. The education leader worked close with the author in this study as she had known the woman for 10 years. The woman was profiled as complex in personality and living at a collective (group). Brought up in a foster family who she did not get along with and she had changed work place many times as she finds problem to fit in and takes up a 'huge space.' She was disruptive, bossy and dominant in the school and talked a lot with many words that she did not understand often trying to involve herself in teacher issues. She had a need to be good (best) and not disabled. The leader felt that the woman had no contact with her own feelings as when others would laugh or cry she would force an enacting of the same. When she cried there was no substance but when she gets angry and screams she has a contact. She has much aggression in her body, always large movements without any nuances. In music she had to be loudest on the drum or similar in other class and has problems within teams. The centre was considering expelling her. I worked on developing exercises associated to awareness of her 'self', her social skills and self-esteem alongside confidence and dynamics.

During the 2 weeks where 3 sessions a day took place she was not allowed in class. We began with a questionnaire (compiled through online research and teacher input, supplemented by input from Centre for Rehabilitation of Brain Injury, Copenhagen) which she completed while being videoed attached to a biofeedback stress indicator (GSR[i] attachment). Each day had variety in which exercises were undertaken but never two the same in one day. The videoed sessions resulted with commented marked progress by staff and other users to her change in behaviour. At the end of the two weeks I took her out in the Centre's van to the city blindfolded and with headset and microphone into which she was to speak of what she perceived as the outside sound being, for example a pedestrian crossing alarm, a person with a shopping trolley, a dog, etc. This was recorded by an authorized concealed video camera in the van. The resulting 30 minute commentary and others from the sessions was played back to her in the company of the leaders and author on the last day. She was asked to comment on the sessions and especially on how she felt about what she saw and how she remembered the sessions. This was a key aspect of the work as she made comments that the leader was in awe about. Her sense of achievement from being able to create with the given audiovisual feedbacks was readily apparent. This seeing and hearing of herself was a powerful tool combined with the sessions. Her 'outside in' reflection from what she saw on the videos seemed to have a positive effect on her in the immediate period at the education when she went back to the classes exhibited by a diminishing of her disruptive effect on fellow users and a quieter disposition. At the start of the two weeks an e-mail account to researcher two was set up and she was shown how to use it as although limited she had verbal and writing skill. She knew him from earlier and she was positive about the idea. This was so as to allow me to have time to prepare the next session, to write up the results and to review and archive via layer analysis the session videos. Response was daily and her diary of events in the sessions told how she felt about it all. The leaders were amazed at some of her writing and even after the sessions had finished and she returned to class she continued writing mails to colleagues and this was a new expressive outlet for her.

5.2 Lund results

Apart from some problems with certain aspects of the questionnaire (translation from English to Swedish and some of the personal questions asked) but following an interview with the user, the leader cleared up the misunderstandings and explained to me the problems. The leader was extremely pleased with the study.

Figure 3. *Movement data captured from the user's brush strokes when she was painting triggers synchronized images and sounds that was relative to the movement. A smooth sound inspired painting with care and smooth strokes - this in turn made the music smoother. Correspondingly, aggressive sound resulted in a disjointed and violent painting action. Both quad images show through facial features and posture an immersive aesthetic resonant experience which at one point resulted in her initiative to paint with two brushes.*

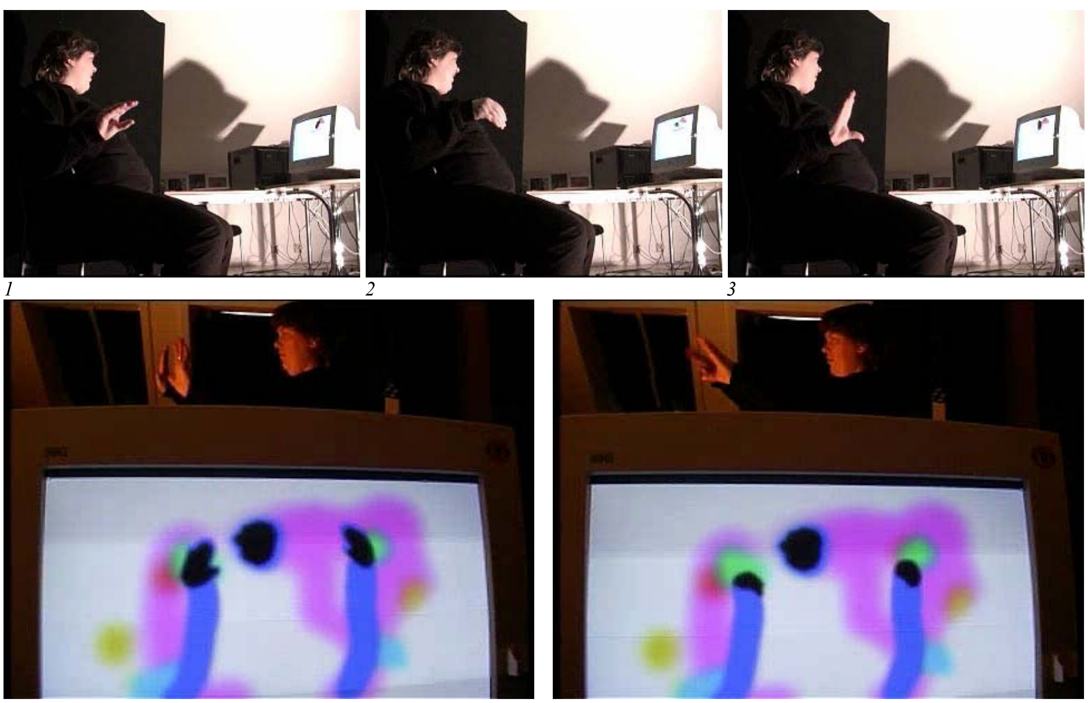


Figure 4. *Hand movement painting being followed on monitor (1, 2, & 3): Lower - similar but she faces the large screen, we see monitor results where the velocity acts as nuance training. Her task was to paint a shape keeping the same velocity which results in the same colour. A deviation in velocity resulted in a change of colour which meant a restart.*

An experiment in the use of the feedback as primary instruction rather than the body part as primary was in two forms of instruction (1) *"When you move your hand in this way the colour will go from dark to light green.,"* = body part as primary (2) "*To make the colour go from dark to light green you will have to move your hand in this way.,*" = feedback as primary. This was to influence the user to think out of the body and in so doing approach the Virtual Interactive Space (VIS) that is inherent between the technology and the user in Aesthetic Resonant Environments, (2) proving more effective with the user mind-set.

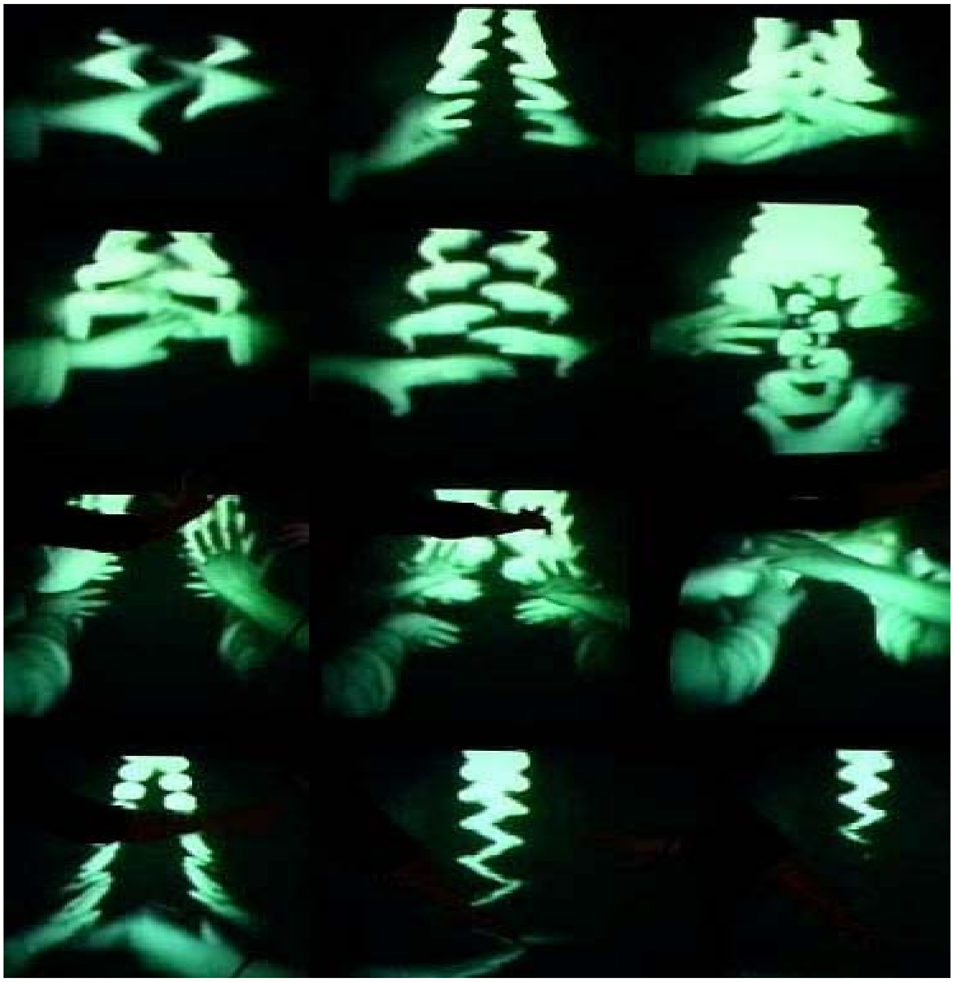


Figure 5. *Video feedback to research potential use as an Eyesweb new algorithm; the upper six images are of our single study alone and the lower six screen shots are from an interaction with a member of the staff. It was wonderful to observe our single case study laughing out very loud as she explored, created and experimented. Her awareness to her hands was total. Interaction with another person was also encouraging.*

5.3 Landskrona

Landskrona (P7) users were six children between 10 – 19 years of age with severe brain damage and additional disabilities. The children are all restricted to use a wheelchair and they have no speech communication. They additionally have mobility and attention problems. Every session was video filmed with MCA to catch the visuals and sounds as well as the facial and the bodily expression of the children. These were then edited and analyzed according to the project goals (inc. LA) and then burned on CD-Rom. After the Personics problems all sessions used the Lund Eyesweb algorithm (as fig. 4) and all users reacted in a positive but different way. This could depend on them having different preferences regarding visual or sound stimuli or both in combination. The few times they have not responded in a positive way could be explained by physical reasons like bad colds, lack of sleeping during the previous night, having had severe or multiple epileptic seizures during the night or day. The epilepsy has never occurred during the sessions.

The sessions resulted in cases with positive reactions of higher motivation in schoolwork, and an increased feeling of well being and pleasure after the sessions. Also it seemed the pupils are further stimulated to see their own real image as a little part of the large visual on the screen which also makes it easier to video both what happens on the screen and catching their facial and body expressions.

Table 1. *An overview of the Landskrona user group of children.*

Ref	Child info.	Focused project work to empower:
A1	Male 19, severe mental retardation + cerebral pares with severe disability + cerebral atrophy + epilepsy. Low awakening grade sometimes makes him sleep for days.	• Awakening grade • Mobility • Expression
A2	Female 10, significant microcephal mental disability + slight deviation in her contact behaviour + walks without support though very spread legged + very hard to motivate regarding exercises of a cognitive nature. A lot of sounds.	• Body awareness • Focusing • Cause +effect • Expression and communication
F1	Female 9, brain damage which severely affects her speech + motor skills + very spastic + cramps. Can take instructions + trains on BLISS method.	• Verbal activity • Coordination/Body awareness • Expressivity
M1	Male 14, brain damage + cerebral pares type spastic tetraplegia + sight impairment + epilepsy + severely disabled + severe mental retardation. Emotionally reacts.	• Motivational aspect • Movement • Communication aspect
M2	Female 18, microcephal + severe brain damage + cerebral pares. No mobility or meaningful motor skill. Severe curvature of the spine + spastic tetraplegia + severe mental retardation + limited contact ability + strong tendency to constipation which makes it important to stimulate her own movements.	• Motor skill • Body awareness • Understanding of cause and effect • Expressivity
P1	Male 6, microcephal mental disability + cerebral pares + explicit spasticity + epilepsy. Alert + contact skill + no language + no mobility.	• Motor skills • Body awareness • Focusing

6. CONCLUSIONS

The biggest success from the project was the creation in Sweden of the Eyesweb painting algorithm reported in Camurri (*et al.* 2003). In both Swedish locations a large 2.5 (w) x 2 (h) meter projection screen was used and this is optimal for one-to-one relationship and sense of presence. A screen-mounted sound system is also used. We observed how free movement that resulted in causal manipulation of interactive responsive sound and visual feedback was able to achieve a state of aesthetic resonance. The creative content was observed as being a motivational tool for the therapy. The data parameter information that enabled this creativity can be a quantifiable entity so as to support and verify results. However the planned data bases of libraries and archive means was not accomplished because of timeframe. The audiovisual interactive content: sonic navigation, visual empathy, sonic tactility, and capture of expressivity was achieved (to a point) through the relational sounds and shapes from our UK partner and the software algorithms. Adaptability of the system was through the graphical user interface sliders but more time was required to build libraries to complement the Eyesweb libraries. The modelling and capture of expressivity however was limited by the timeframe and we did not work with physical models which would have been optimal. The commercialisation of the technical innovations integration into the Personics product was curtailed by the consortium. Personics have since commercialised their own independent product.

As a result of the project we have established new qualitative and quantitative user evaluation methodologies most appropriate to our wide user base. This will be exploited more fully in the next project with the creation of much needed new testing methods for process analysis and user categorization. The research which the Swedish partners conducted is ongoing by the co-authors where we are using the Soundbeam ultrasound sensor in conjunction with the Eyesweb camera system. The Soundbeam is a mature technology that was used in the prior research and is acknowledged as an industry leader. It is more stable than the Personics infrared sensor and unlike the Personics sensor it gives a predictable intuitive response to the user. Our Italian software partners continue to develop Eyesweb. The project proposal identified the Minimum Data Set / Resident Assessment Instrument (MDS/RAI) as a proven assessment tool for measuring the impact of CARE HERE interventions with older people and with Parkinson's patients and to explore its possible applicability to the other client groups in the study, notably younger people with severe physical and learning disabilities. The tool was shared with the partners and tested for suitability. This was rejected by the Swedish researchers as unsuitable for their user groups.

Acknowledgements: The authors would like to thank all the *non-Personics* consortium personnel who worked so hard to realise the project and especially those involved in the Swedish research and the developed network that assisted us in our endeavours in the South West corner of Sweden.

7. REFERENCES

A L Brooks (1999) Virtual Interactive Space (V.I.S.) as a Movement Capture Interface Tool Giving Multimedia Feedback for Treatment and Analysis, *Proc. 13th Int. Congress of World Confederation for Physical Therapy*, Yokohama, Japan, pp. 66.

A L Brooks and S Hasselblad (2001) Non wearable HCI sensor system control of 3D images and other feedback through movement capture of an eyelid flicker, a head gesture or even full body, *Workshop, Proc. Advances in Human-Computer Interaction*, Patras, Greece, pp. 9.

A L Brooks, S Hasselblad, A Camurri, N Canagarajah (2002) Interaction with shapes and sounds as a therapy for special needs and rehabilitation, *Proc. 4th Int. Conference On Disability, Virtual Reality, and Associated Technologies,* Veszprém, Hungary, pp. 205-212.

A L Brooks (2004a) HUMANICS 1: A study to create a home based telehealth product to supplement acquired brain injury therapy, *Proc. 5th Int. Conf. On Disability, Virtual Reality, and Associated Technologies,* Oxford, England, *In press*.

A L Brooks (2004b) Soundscapes, In *Inhabited Information Spaces-Living with Your Data,* D N Snowdon, E F Churchill and E Frécon (Eds.), Springer-Verlag London, pp. 89-100.

A L Brooks (2004d) 'Interactive Painting' – An evolving study to facilitate reduced exclusion from music concerts for the Deaf community, *Proc. 5th Int. Conf. On Disability, Virtual Reality, and Associated Technologies,* Oxford, England, *In press*.

A Camurri, B Mazzarino, G Volpe, P Morasso, F Priano and C Re (2003) Application of multimedia techniques in the physical rehabilitation of Parkinson's patients, J. Vis. Comp. Animat, 14, pp. 269-279.

P Ellis (1996) Layered analysis: A video-based qualitative research tool to support the development of a new approach for children with special needs. In; the *Bulletin for the Council for Research in Music Education,* University of Illinois at Urbana-Champaign, USA, **130**, 65-74.

R Kizony, N Katz, and P Weiss (2003) Adapting an immersive virtual reality system for rehabilitation, *J. Visual. Comput. Animat.*, **14,** pp. 261-268.

J Lanier (1992), *Virtual Reality and Persons with Disabilities conf. kn. address*, www.csun.edu/cod/vr92.html.

O Martin, B Julian, L Boissieux, J gascuel and C Prablanc (2003), Evaluating online control of goal-directed arm movement while standing in virtual visual environment, *J. Visual. Comput. Animat.*, **14,** pp.253-260.

C Paccheti, R Aglieri, F Mancini, E Martignoni and G Nappi (1998) Active music therapy and Parkinson's disease: methods, *Functional Neurology*, **13**, 57-67.

C Paccheti, R Aglieri, F Mancini, E Martignoni and G Nappi (2000) Active music therapy and Parkinson's disease: an integrative method for motor and emotional rehabilitation, *Psychosomatic Medicine*, **62**, pp.386-393.

D Rokeby (1990), Very Nervous System (VNS): 'Canadier omsætter video til musik,' *Ingenioren*, **20**, pp. 30.

A Rutten, S Cobb, H Neale, S Kerr, A Leonard, S Parsons and P Mitchell (2003), The AS interactive project: single-user and collaborative virtual environments for people with high-functioning autistic spectrum disorders, *J. Visual. Comput. Animat.*, **14**, pp. 233-241.

H Yano, K Kasai, H Saitou and H Iwata (2003), Development of a gait rehabilitation system using a locomotion interface, *J. Visual. Comput. Animat.*, **14**, pp. 243-252.

[1] Pressing links at - *www.psych.unimelb.edu.au/staff/pressing.html*

[2] National and International - *www.bris.ac.uk/caress* - *www.cfh.ku.dk/6_forskning/Humanics*

---

[i] Galvanic Skin Response – a biofeedback technique used in lie detection.

AudioBattleShip: blind learners cognition through sound

J H Sánchez

Department of Computer Science, University of Chile
Blanco Encalada 2120, Santiago, CHILE

jsanchez@dcc.uchile.cl

ABSTRACT

Recent literature provides initial evidence that sound can be used for cognitive development purposes in blind children. In this paper we present the design, development, and usability testing of AudioBattleShip, a sound-based interactive environment for blind children. AudioBattleShip is an interactive version of the board Battleship game, providing different interfaces for both sighted and blind people. The interface is based on spatialized sound as a way of navigating and exploring through the environment. The application was developed upon a framework that supports the development of distributed heterogeneous applications by synchronizing only some common objects, thus allowing the easy development of interactive applications with very different interfaces. AudioBattleship was tested for cognitive tasks with blind children, evidencing that it can help to develop and rehearse abstract memory through spatial reference, spatial abstraction through concrete representations, haptic perception through constructing mental images of the virtual space, and cognitive integration of both spatial and haptic references.

1. INTRODUCTION

Diverse audio-based applications for blind children have being implemented in the last few years (Baldis, 2001; McCrindle and Symons, 2000; Mereu and Kazman, 1996; Sánchez, 1999, 2001). Many of them focus on the development of 3D audio interfaces to map the entire surrounding space as proof of concept concerning audio-based interfaces. Very few studies centre on the effects and impact of sound on developing cognitive skills of blind people by evaluating the cognitive usability of these applications. No previous work has centred on using spatialized sound to develop abstract memory, haptic perception, and the integration of spatial and haptic references by blind learners. We also have no reference of previous work for engaging blind and sighted people in interactive and collaborative games.

Developing interactive systems for integrating people with different capabilities poses a higher challenge because of the synchronization of different interfaces for each type of user. Interfaces for sighted users normally consist of computer graphical user displays and some sound for the output. The keyboard and mouse are the most common devices for input. Interfaces for blind users rely principally on auditory information for output, keyboards, and other haptic devices for input (Tan, 2000; Sjostrom, 2001). Sometimes joysticks are used for both input (movement) and output, as they pose some resistance in order to simulate the contour objects.

This research study presents the design, development, and usability testing of AudioBattleShip, a sound-based interactive environment for blind children (Sánchez et al., 2003). This system is a similar version of the traditional battleship game for sighted people but including both graphical interfaces for sighted users and audio-based interfaces for blind people. AudioBattleShip provides a gaming environment to enhance group interaction, abstract memory, spatial abstraction, and haptic perception in blind learners. We also introduce a platform which supports programming of distributed applications, especially for synchronizing applications having different interfaces.

For designing and developing information-equivalent interfaces (Veijalaien and Gross, 2001) for sighted and blind people we followed a similar process such as the one described in (Baloian, Luther & Sánchez, 2002) for developing interfaces for people with disabilities.

Finally, a full usability study has been implemented to evaluate the cognitive impact when blind children interact with AudioBattleShip.

2. THE AUDIOBATTLESHIP DESIGN

2.1 Playing modes and phases

As the aim of this game is to engage blind people in a interactive environment to help to develop cognitive skills we decided to provide various playing modes. AudioBattleShip can be played blind to blind, blind to sighted, and blind to computer. Blind to blind mode presents the same interfaces to both players. Blind to sighted mode provides a variety of tools to the blind learner to minimize the disadvantages compared to sighted learners that can have in any moment snapshots of the state of actions. Blind to computer mode gives intrinsic advantages to the computer over the blind learner because of computer memory and probability calculations, but this disadvantage can be diminished by limiting the algorithm and creating different levels of complexity.

AudioBattleShip has three phases (see Figure 1):

Ship positioning phase: the player chooses the position of ships taken from a predefined set on the battlefield. A matrix represents a battlefield where ships can be placed over a column and a row by covering different numbers of spaces according to the type of ship.

Creating and joining session phase: After placing the ships, a player can choose between creating a new session of the game, joining an existing one, and playing against the computer. If a new session is created, another player can join it. To join an existing session in the Internet the player has to know the session's name and the host address.

Shooting phase: By taking turns both players can manipulate the matrix with the board's state by identifying a cell within the matrix representing the contender's battlefield for dropping a bomb on that place. The system responds whether a contender's ship was hit or not.

Figure 1. *Phases of the AudioBattleShip game*

2.2 Interface design

Baloian, Luther & Sánchez (2002) propose a model for developing educational software for people with different disabilities by using diverse interfaces. This model can be briefly summarized in the following steps: 1. To develop the computational model of the real world; 2. To model the input/output needs for the system (including feedback); and 3. To develop interfaces by "projecting" the model over a range of user's interaction capabilities.

In this case, we modelled AudioBattleShip by mapping the two players, each with a matrix representing the space where their own ships are positioned. A player's matrix will also register the information about where the contender has dropped a bomb. A token represents which player is in turn for dropping a bomb in the contender's field.

We created two different applications, one for blind and one for sighted users. Each application has two interfaces, one for the ship positioning phase and the other for the shooting phase. The ship positioning interface for sighted people consists of a window containing a grid representing the battlefield. Ship positioning takes place by defining a position and a direction (up, down, left, right) with the mouse. The shooting phase interface consists of a window displaying two matrices, one for the battlefield where the player's own ships are positioned and the other representing the contender's field, showing the places where the player has already drooped bombs and the outcome of this operation. A text area was provided to inform about the course of the game.

The application for blind users uses a tablet as input device (see Figure 2a). The tablet can map the entire screen and by using a pen-based pointing device diverse mouse events can be triggered. A grid was built over the tablet in order to represent the matrix of the battlefield and some additional "help buttons" for triggering actions. During the ship positioning phase the player has to point to a certain place on the grid and indicate the direction by moving the pointing stick to a cell up, down, left, or right. During the shooting phase the player has just to double-click over a certain cell of the grid (see Figure 2b). Sound feedback is provided to inform about a specific spatial location on the board and the occurrence of certain actions, for example, the resulting outcome of dropping a bomb in a cell of the contender's battlefield. Help buttons assist the player in remembering the places where an enemy's ship was hit or not.

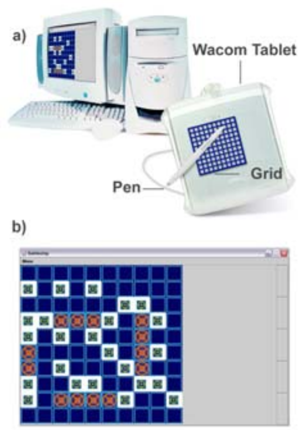


Figure 2. *Input devices (a) and shooting phase (b) for blind players*

3. A PLATFORM FOR REPLICATED COLLABORATIVE APPLICATIONS

Building collaborative software, a developer has to decide a mechanism for making a distributed application. Basically, the developer has to decide between two paradigms. One of them is to select a framework with a centralized server who is capable of serving several clients. In this case, the information needed by all clients is first sent to the server and then somehow broadcasted to the clients. In these scenarios, e.g. NetMeeting (www.microsoft.com/windows/netmeeting), it is not possible to work with the clients any longer if the server is for some reason not accessible any more. The second paradigm is a replicated architecture. Here all the necessary data for clients is replicated in the client applications. The advantage of this kind of framework is that clients are still operable even if they lose their network connection.

The MatchMaker TNG framework (Trewissen et al., 2000), which is based on RMI, combines both paradigms for making collaborative software. First of all it has a centralized server, which manages to send the needed data to the clients, and it also replicates the whole internal data structure of the application. By replicating the data structure it is possible for each application to have a comparative synchronized status. Within this server application the data structure is arranged as a tree called the "synchronization tree". It is not necessary for every client to listen to the whole tree but clients may also listen to certain sub-trees. Most likely, the synchronization tree will reflect the internal data structure of the application as shown in Figure 3.

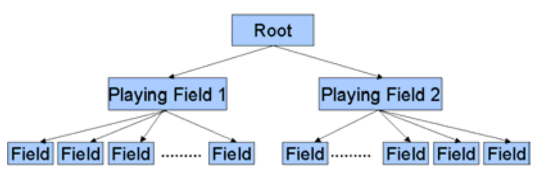


Figure 3. *The synchronization tree for AudioBattleShip*

Several frameworks for synchronization use event-based technologies (Jansen et al., 2001), but MatchMaker TNG replicates according to the Model-View-Controller paradigm, by implementing a so-called MatchMaker TNG model for every object to be replicated. This MatchMaker TNG model consists of the object model but may also have information about the view. By using these models the possibility of interpreting them differently in diverse applications is high. For example a model may be fully interpreted by an application that is running on a PC with high performance, but some pieces of information may not be considered if the same model is interpreted by an application running on a PDA.

Another example is the use of the model in AudoBattleShip. Here it is possible to integrate in the model information needed for people with disabilities. This could be ignored by an application that is running for people without disabilities.

The models are propagated to the clients using the common event architecture of Java. The server has four basic operations to inform the clients about changes in the synchronization tree when a new node is created, deleted, changed, or activated. Every MatchMaker TNG client has to implement a special method for each of these four operations.

3.1 The implementation of AudioBattleShip

Since the interfaces of both applications are very different the most convenient way to synchronize them is at the level of the data model. AudioBattleShip has objects that store the information about players, the state of the game (turns), and the coordinates of a given point with the corresponding state (ship, water, destroyed ship, shooting to water). By using these objects a synchronization tree is created with nodes that store information about players that also contains objects corresponding with the size of the board.

According to this, the MatchMaker synchronizing tree has the turn-taking token at the root of the tree. The information about the two players is stored as two child nodes of the root. Each cell representing the battlefield of a player is stored as a separate child node from the player's node (see Figure 3). This has the advantage that only the information of one cell will be transmitted to the coupled application if just the information changes, and not the whole matrix. Synchronizing applications using sound have also an additional problem: the sound playing is usually implemented as a separate thread in most programming languages (and Java is not the exception). In many cases, we must be sure that some auditory information has been completely heard by the user before continuing with the program execution (which may mean the playing of another sound clip). To accomplish this, Java Media Framework was used to synchronize multimedia events. AudioBattleShip contains a directory of sounds that can be easily modified to change language and sounds.

Figure 4. *On site interaction with AudioBattleShip between two blind learners*

4. USABILITY TESTING

AudioBattleShip was usability tested both during and after software implementation. Three usability experiences were designed. Two of them were centred on the usability of the interface provided and the last testing was concerned with the cognitive impact of using AudioBattleShip.

4.1 Interface usability

Interfaces of AudioBattleShip were evaluated by blind learners. The first experience consisted of testing different software modules with four blind learners that regularly attend the "Santa Lucia" School in Santiago, Chile. After finishing a software module, it was tested by learners. They interacted with the software, made comments, and answered questions. Thinking aloud methods were also implemented. All of these procedures helped designers to redesign and make it closer to the mental model of the blind children.

The first prototypes were fully tested. The interface design and functionality were analyzed. Four blind learners from the same school participated in the study. They played the game and answered questions during four sessions. Two people from the research team observed them, asked questions, and registered their comments and behaviours in all sessions. Most questions were related to audio interfaces and the mapping and use of input devices such as the tablet. We tested how well the concrete matrix of the traditional board Battleship game was mapped into the interactive environment. The mapping of the screen matrix by the

Proc. 5th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, UK, 2004

tablet with a rough grid on it using a synthetic material was also evaluated. We then tested the usability of some interface elements such as sounds (intensity and quality), volume, and feedback to the user actions.

The tasks followed by learners were the following: 1. To explore AudioBattleShip, 2. To understand the structure, rules, and adaptations made, and 3. To understand the differences between the concrete Battleship board and the digital version of the game. Users also learned how to play the game by using both the keyboard and tablet. They fully explored the pros and cons of using a tablet, its functions and analogies with the AudioBattleShip interface. We looked for audio stimuli association to game actions and the understanding of cause-effect action relationships when interacting with the software (see Figure 5).

As a result of this preliminary testing, learners enjoyed interacting with AudioBattleShip. They helped to redesign the software by making insightful and meaningful contributions about sound synchronicity, sound overlapping, sound help, colour contrast, size of the cursor, position identification, and tablet mapping.

Figure 5. *Interface usability testing by blind learners*

4.2 Cognitive impact

We tested the cognitive impact of AudioBattleShip with five learners during six sessions of ninety minutes each. All sessions were captured and registered using a video digital camera. The comments made and answers to specific questions were also registered. Two special education teachers observed blind learners doing tasks with AudioBattleShip using the tablets and filled out observation forms specifically designed for this purpose. We developed a paper form for each cognitive task containing several questions, an answer scale, and open observations.

The sequence of actions followed by learners included: 1. To navigate AudioBattleShip, 2. To identify different audio stimuli, 3. To explore both the concrete board of the traditional Battleship game and the Wacon tablet, 4. To play AudioBattleShip, and finally, 5. To make concrete representation models (clay and Lego) of navigating spaces and position moves (see Figure 7). They interacted with their partners and then represented in a concrete board the moves of the adversary. One player described the coordinates and others localized the pieces and marked them on the board. They also searched the ships of the adversary by using a strategy previously designed (see Figures 4 and 6).

We looked for the impact of using AudioBattleShip on abstract memory through spatial references with sound, spatial abstraction through the concrete representation of the virtual environment, and haptic perception (tactile and kinaesthetic) by generating mental images of the navigated space. They also integrated spatial references through sound and haptic references through the tactile manipulation to construct mental images of the navigated virtual environment.

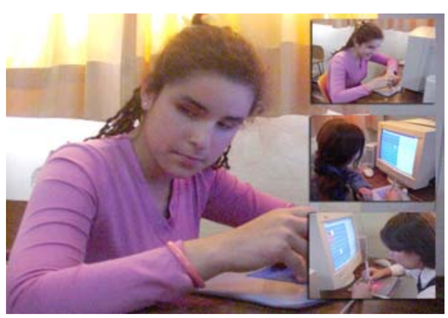


Figure 6. *Evaluating the cognitive impact of using AudioBattleShip*

Figure 7. *Concrete representation models of the navigating space and position moves*

Task one was concerned with integrating spatial references through sound and haptic references through tactile manipulation to construct mental images of the navigated virtual environment. We also wanted to check whether or not by using AudioBattleship learners can perceive the virtual space through haptic perception by generating mental representations. Most learners could clearly differentiate each audio stimulus. Most of them could make an excellent analogy between AudioBattleShip and the concrete board of the traditional Battleship game (80%). The same percentage of learners could make an analogy between the Wacon tablet and the virtual environment of the adversary. All learners did not need a concrete board to orientate their movements and navigation throughout AudioBattleShip (see Figure 8). Some learners performed excellent against an adversary (40%) and others did a good performance (60%). The same result was obtained when played against the computer. This means that learners could design a strategy to solve problems and defeat the adversary.

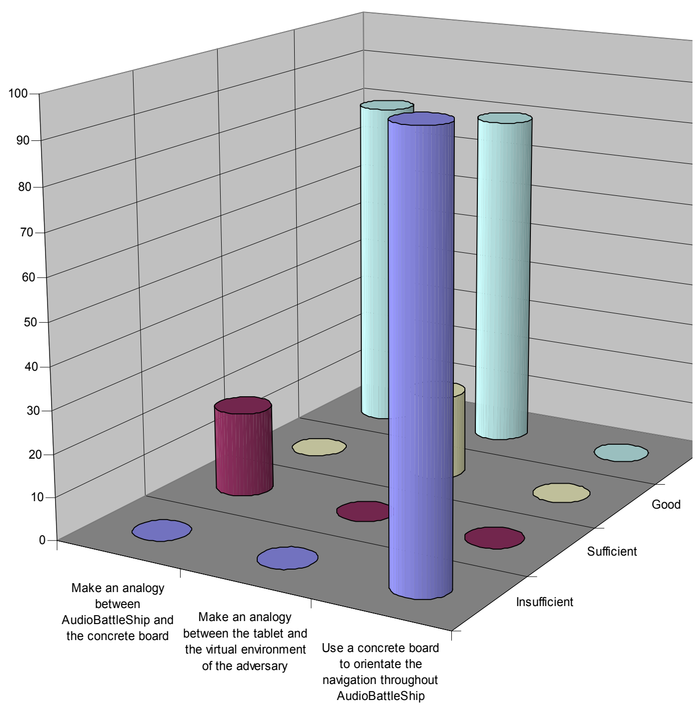


Figure 8. *Mental images and haptic perception through using AudioBattleShip*

Task two measured abstract memory through spatial references with sound and spatial abstraction through the concrete representation of the virtual environment. Almost 90% of learners could completely represent in a concrete board the moves of the adversary. None of them needed to use a concrete board to orientate their navigation throughout AudioBattleShip (see Figure 9).

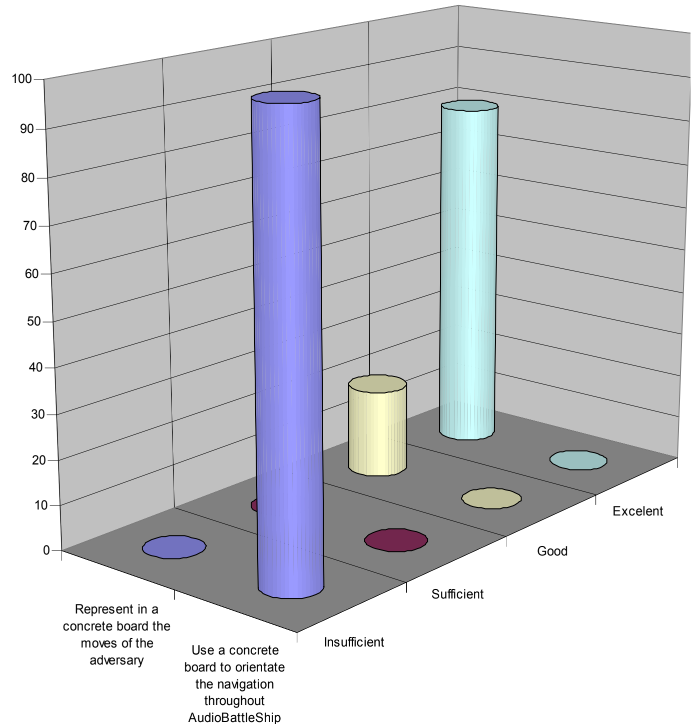


Figure 9. *Abstract memory and spatial abstraction through AudioBattleShip*

Task two also included a specific activity related to detect how well learners could develop a strategy to overcome the adversary. Most of them could design and implement a strategy to win the game (80%). All of them could orientate easily within the matrix implemented on the Wacon tablet.

5. FINAL DISCUSSION

This study was aimed at designing and evaluating a sound-based virtual environment to develop and rehearse abstract memory through spatial reference, spatial abstraction through concrete representations, haptic perception through constructing mental images of the virtual space, and cognitive integration of both spatial and haptic references. Our data indicate that blind learners using AudioBattleShip mediated with cognitive tasks and concrete materials can help to develop and exercise mental images of the space, haptic perception, abstract memory, and spatial abstraction.

There were diverse forms of interaction among blind learners. Some collaborative abilities can be enhanced by using interactive sound-based applications such as AudioBattleShip. We need to know in more detail what type of skills can be stimulated and ways of improving them through the interaction with sound. This is especially important for blind learners since they are accustomed to doing individual work with little social interaction when using digital devices. We envision constructing new ways and scenarios for collaboration between blinds as well as between blinds and sighted learners by using digital audio interfaces such as the one introduced here.

Learners enjoyed the interaction with AudioBattleShip. This can be explained because this interaction allowed them to compete and measure their competences and abilities. This experience also had an affective side. They did improve their self-esteem and self-value to compete defeat other learners with and without vision. Learners also showed more interest to play with a blind partner instead of playing against the computer. They think interacting with computers did not challenge and give the same feedback in comparison to playing with a peer. They were motivated to play and interact with sighted learners because it was a way to show they can make the same performance as sighted if minimal conditions are met at the beginning.

The use of Wacon tablets as interaction devices for AudioBattleShip was a critical strength. They allowed blind learners to interact with different interface haptic devices and demonstrated to be highly appropriate for this type of software.

A recent retest study (Sánchez, Baloian, and Hassler, 2004) is giving us more detailed insights and knowledge about how cognition and collaboration can be enhanced by the interaction with sound. We are contrasting this quantitative data to evaluate more fully the real contribution of audio-based interfaces in blind children cognition and collaboration.

Finally, the observations made by learners, their comments, and meaningful ideas were critical to improve AudioBattleShip. This validates one of our premises in a sense that interactive software environments based on sound should be designed with and for blind learners in order to map their needs, interests, and mental models.

Acknowledgment: This report was funded by the Chilean National Fund of Science and Technology, Fondecyt, Project 1030158.

6. REFERENCES

C Sjostrom (2001), Using haptics in computer interfaces for blind people. *Proceeding of the ACM CHI ´01*, Vol 3, 1, pp. 245-246.

F Tewissen, N Baloian, H Hoppe, and E Reimberg (2000), "MatchMaker" – Synchronising objects in replicated software architectures. In *Proceedings of CRIWG 2000 (Sixth International Workshop on Groupware*), Madeira, Portugal, October 2000, Los Alamitos: IEEE Press. pp. 60-67.

H Tan, (2000), Haptic interfaces. *Communications of the ACM*, 43(3), pp. 40-41.

J Baldis (2001), Effects of spatial audio on memory, comprehension, and preference during desktop conferences. *Proceeding of the ACM CHI ´01*, Vol 3, 1, pp. 166-173.

J Sánchez (1999), Interactive 3D sound hyperstories for blind children. *Proceedings of ACM-CHI 99*, pp. 318-325.

J Sánchez (2001), Interactive virtual acoustic environments for blind children. *Proceedings of ACM CHI ´2001*, Seattle, Washington, pp. 23-25.

J Sánchez, N Baloian, & T Hassler (2004). Blind to Sighted Children Interaction through Collaborative Environments, X International Workshop on Groupware, CRIWG 2004, September 5 al 9, San Carlos - Costa Rica (will be also published in an issue of *Lecture Notes in Computer Science, LNCS*, Springer-Verlag).

J Sánchez, N Baloian, T, Hassler, and U Hoppe (2003). AudioBattleship: Blind Learners Collaboration Through Sound. *Proceedings of ACM CHI 2003*, pp. 798-799. Also published by *SIGCHI publications, ACM Press*. Fort Lauderdale, Florida, April 5-10.

J Veijalainen and T Gross (2001), Mobile Wireless Interfaces; in Search for the Limits. In: *Developing an Infrastructure for Mobile and Wireless Systems*; König-Ries, B.,Makki, K., Makki, S.A.M., Pissinou, N., Scheuermann, P., (Eds.) NSF Workshop IMWS 2001, Scottsdale, AZ, October 15, 2001, Revised Papers. Springer Verlag LNCS Vol. 2538, Dec. 2002, pp. 153-163.

M Jansen, N Pinkwart and F Tewissen (2001), MatchMaker - Flexible Synchronisation von Java-Anwendungen. In Klinkenberg, R., Rüping, S., Fick, A., Henze, N., Herzog, C., Molitor, R. & Schröder, O. (eds.): *LLWA 01 - Tagungsband der GI-Workshopwoche "Lernen-Lehren-Wissen-Adaptivität",* October 2001. Forschungsbericht 763. Universität Dortmund, Germany.

N Baloian, W Luther and J Sánchez (2002), Modeling Educational Software for People with Disabilities: Theory and Practice. *Proceeding of the ASSETS 2002 Conference*, 8 - 10 July 2002, Edinburgh, pp. 111-117.

R McCrindle and D Symons (2000), Audio space invaders. *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies*, pp. 59-65.

S Mereu and R Kazman (1996), Audio enhanced 3D interfaces for visually impaired users. *Proceedings of CHI ‘ 96, ACM Press*.

ICDVRAT 2004

Session VII. Brain Injury and Rehabilitation

Chair: Craig Carignan

Employing a virtual environment in postural research and rehabilitation to reveal the impact of visual information

E A Keshner[1,2], R V Kenyon[3], Y Dhaher[1,2] and J W Streepey[1]

[1]Sensory Motor Performance Program, Rehabilitation Institute of Chicago,
345 East Superior Street, Chicago, IL 60611, USA

[2]Dept. of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University,
303 East Chicago Avenue, Chicago, IL 60611, USA

[3]Department of Computer Science, University of Illinois at Chicago
851 South Morgan, Chicago, IL 60607, USA

[1]*eak@northwestern.edu*, [2]*kenyon@uic.edu*, [3]*y-dhaher@northwestern.edu*, [4]*jstreepy@northwestern.edu*

[1]*http://sulu.smpp.northwestern.edu/~keshner,* [2]*http://sulu.smpp.northwestern.edu/dhaher,*
[3]*http://www.cs.uic.edu/~kenyon*

ABSTRACT

We have united an immersive virtual environment with support surface motion to record biomechanical and physiological responses to combined visual, vestibular, and proprioceptive inputs. We have examined age-related differences during peripheral visual field motion and with a focal image projected on to the moving virtual scene. Our data suggest that the postural response is modulated by all existing sensory signals in a non-additive fashion. An individual's perception of the sensory structure appears to be a significant component of the postural response in these protocols. We will discuss the implications of these results to clinical interventions for balance disorders.

1. INTRODUCTION

In an effort to understand how the central nervous system (CNS) controls motor performance in humans, scientists have endeavoured to reduce the complexity of the system to a single pathway of control and a single measurable output. However, data from this approach are not robust because the CNS is very plastic and will alter its responses and control parameters over time and with specific parameters of a task. Thus, if the task is to walk from one location to the next with the eyes closed, we are not necessarily learning about the role of vision in locomotion, but about how the CNS will adjust its responses when vision is not available. In fact, the visual system has been downplayed as a contributor to postural control because it has a longer reaction time than somatosensory inputs and because when removed, either by closing the eyes or by darkening the room, no significant changes in automatic postural reactions have been observed (Buchanan and Horak, 1999; Keshner et al, 1987; Nashner and Berthoz, 1978). But in natural rather than experimentally controlled environments, visual signals may have a greater influence on postural orientation than expected from laboratory studies. Using a virtual environment (VE), we tested such a situation (Keshner and Kenyon, 2000). Subjects were asked to walk while viewing a virtual representation of a room that rotated at a constant velocity about the visual axis of the subject. The roll motion of the room was uncorrelated with the parameters of their locomotion. We observed that subjects either altered the organization of their locomotion pattern or lost their balance while walking. Thus, when findings from the laboratory are applied to therapeutic interventions, the intervention may not be appropriate for all circumstances and will not fully meet the needs of the patient.

The impact of visual information on movement planning is particularly relevant to studies of postural control because maintaining posture is a multimodal process. During movement we are simultaneously exposed to visual inputs, information from the ground, and somatosensory signals from our own body that may be transient or sustained and that occur at multiple frequencies and in multiple directions with respect to our motion. The performer may choose to suppress responses to irrelevant or conflicting sensory information, thereby simplifying the control process. However, if conflicting inputs are not suppressed, but can instead

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

influence and modify the weighting of all other sensory inputs, then response parameters must accommodate even inappropriate sensory signals. The consequence of this, particularly in the elderly (Peterka et al, 1990), is that the performer may not detect, or may not plan a response that will match, the most imperative stimulus parameters and a fall will occur.

We have attempted to resolve this insufficiency by developing the Virtual Environment and Postural Orientation (**VEPO**) laboratory. Although this laboratory continues to evolve (Keshner and Kenyon, 2004), all of our studies combine biomechanical and physiological measurements with an experimentally controlled immersive wide field-of-view visual environment. We present a series of studies here in which we examined the relative weighting of visual and physical stimuli on the postural response in both healthy young and elderly adults. A subsequent study in which we compare the effects of mental calculation to those of searching for a visual target presented against a moving visual field is also presented. We presented subjects with sinusoidal inputs so that predictive mechanisms were also be engaged. Thus we have used the VE to keep all inputs activated and have observed that rather than shifting from one signal to another, subjects incorporated characteristics of all inputs into their segmental responses. Adding the virtual environment to the traditional posture laboratory permits the exploration of more complex motor behaviours, and should expose modifications in motor behavior that take place with impairment or disorders of the system thus supporting the development of appropriate rehabilitation interventions.

2. METHODS

2.1 Development of the Virtual Environment and Postural Orientation Laboratory

In the projection-based VE that we have chosen to employ, the computer generated imagery is back-projected on a screen or wall that is in front of the user much like that in a theatre (Cruz-Neira et al, 1992). We use back-projection instead of front-projection to insure that the subject's body does not cast a shadow on the projected scene. Our laboratory consists of one wall of back projection material measuring 1.2 m x 1.6 m. An Electrohome Marquis 8500 projector (Electrohome Ltd., Kitchener, Ontario CA) throws a full-colour stereo workstation field (1280x1024 stereo) at 120 Hz onto the screen. A dual Xenon processor PC with an nVidia Quadro4 900XGL graphics card creates the imagery projected onto the wall. Field sequential stereo images generated by the PC are separated into right and left eye images using liquid crystal stereo shutter glasses (Crystal Eyes, StereoGraphics Inc., San Rafael CA). These glasses limit the field of view to 90° in the horizontal and 55° in the vertical direction. The correct perspective and stereo projections for the scene are computed using coordinates for the current orientation of the head supplied by an infrared motion analysis system. For this we have used both the Optotrak (NDI, Ontario, Canada) and the Motion Analysis (Motion Analysis Corp., Santa Rosa CA) systems. Consequently, virtual objects retain their true perspective and position in space regardless of the subjects' movement. The total display system latency measured from the time a subject moves to the time the resulting new stereo image is displayed in the environment is 20-30 ms. The *stereo* update rate of the scene is 60 Hz which is half the rate at which we sample the head data (120 Hz).

Figure 1. *(Left) Subject wearing stereo shutter glasses stands on the posture platform in front of the back projection screen. The virtual image with a projection of the stationary no-calculation task (centre) and the randomly moving calculation task (right).*

2.1.1 Scene Characteristics. The scene consisted of a room containing round columns with patterned rugs and painted ceiling (Fig. 1). The columns were 6.1 m apart and rose 6.1 m off the floor to the ceiling. The rug patterns were texture mapped on the floor and consisted of 10 different patterns. The interior of the room measured 30.5 m wide by 6.1 m high by 30.5 m deep. The subject was placed in the centre of the room between two rows of columns. Since the sled was 64.8 cm above the laboratory floor the image of the virtual room was adjusted so that its height matched the sled height (i.e., the virtual floor and the top of the sled were coincident). Beyond the virtual room was a landscape consisting of mountains, meadows, sky and clouds.

The floor was the distance from the subject's eyes to the virtual floor and the nearest column was 4.6 m away. The resolution of the image was 7.4 min of arc per pixel when the subject was 40 cm from the screen. The view from the subjects' position was that objects in the room were both in front of and behind the screen. When the scene moved in fore-aft, objects moved in and out of view depending on their position in the scene.

2.1.2 Experimental Protocols. A linear accelerator (sled) that could be translated in the anterior-posterior direction was controlled by D/A outputs from an on-line PC. The sled was placed 40 cm in front of the screen on which the virtual image was projected (Fig. 1). Both healthy young (20-38 years) and older (58-78 years) adults have been tested in this apparatus. Subjects were exposed to a ±15.7 cm/sec sinusoidal translation of the sled in the anterior-posterior direction at 0.25 Hz (±10 cm excursion). A ± 3.8 m/sec sinusoidal fore-aft motion of the computer generated stereo image was presented at 0.1 Hz (±6.1 m excursion) either separately or combined with the sled motion. Each trial lasted a total of 205 sec.

In a second study, the visual scene was translated in fore-aft at 0.1 Hz without a secondary task (scene only) and the sled was anterior-posterior translated (0.25 Hz) while the subject's eyes were closed (sled only). Then, while the sled and visual field were translating in fore-aft, pairs of numbers were projected directly in front of the subject (stationary) or moved randomly within the visual field (moving) (Fig. 1). Subjects were instructed to press a button whenever a stationary or moving pair of zeros appeared (no-calculation) and to press one of two buttons when the difference between the stationary or moving numbers was equal or not equal to 4 (calculation).

2.1.3 Data Analysis. Three-dimensional kinematic data from the head, trunk, and lower limb were collected at 120 Hz by an infrared motion analysis system. Segmental excursions of the head with respect to the trunk, the trunk with respect to the shank, and the shank with respect to the sled were calculated. Center of mass (COM) of the head, trunk, and shank were also calculated. Power of the response at each stimulus frequency was derived using a 40 sec sliding window following a fast Fourier transform analysis. Differences between the populations and across conditions were calculated using Wilcoxon matched-pairs signed rank tests and Wilcoxon rank sums tests.

3. RESULTS

3.1 Responses to Visual Field Motion in Young Adults

We know that when the support surface (i.e., the sled) moves but the world appears stationary, subjects will either ride along with the motion of the sled or compensate for that motion by moving in the opposite direction (Buchanan and Horak, 1999). This indeed was the result we observed in our subjects. Segmental output was reasonably consistent across the whole period of the trial (Fig. 2) and response frequencies peaked at the frequency of stimulus input. When only the visual scene moves but the support surface remains stationary, subjects can choose to ignore the apparent motion of the visual world because both the vestibular and somatosensory systems would signal an absence of physical motion (Dichgans et al, 1972). Some of our subjects could suppress the response to visual motion, but others responded to it, mostly by moving their trunk and head in the direction of the visual scene as would be expected when there is a visual-vestibular conflict (Dichgans et al, 1972).

When both the sled and visual scene were moving, there was a conflict between the visual and the other sensory systems because of the spatial and temporal discordance in the two stimuli. Now it became more difficult to suppress responses to motion of the visual world and subjects responded in one of two ways. Either they locked their segments together and responded as an inverted pendulum, or they incorporated both the frequency of the sled and the frequency of the scene into their segmental responses, often with an associated increase in segmental response amplitude (Fig. 2). The increased power observed in this condition could not be obtained simply by summing the effects of the trials with single frequency inputs. Power of the response increased around 60-80 sec into the trial, similar to a response we may expect if a subject was experiencing vection - the sensation of self-motion through the visual system (Hettinger et al, 1990). With vection, the optical flow pattern creates a compelling illusion of self-motion that is not corroborated by the inertial forces transmitted through the vestibular sense organs.

We have employed a Principal Component Analysis (PCA) to determine the overall weighting of the inputs from the sled and the scene (Fig. 2). For most subjects, the relative weighting of each input fluctuated across a trial, but some subjects exhibited a strong preference for either the sled or the scene. For example, S1 in Fig. 2 had a stronger, more consistent response to the scene motion than to sled motion. S3 demonstrated the opposite effect. Both subjects weighted the two signals relatively equally when presented with the both at the same time. S2, on the other hand, did not respond strongly to either input either when presented singly or in a combined fashion. Contributions of each body segment to the overall response strategy were observed primarily in the trunk and shank.

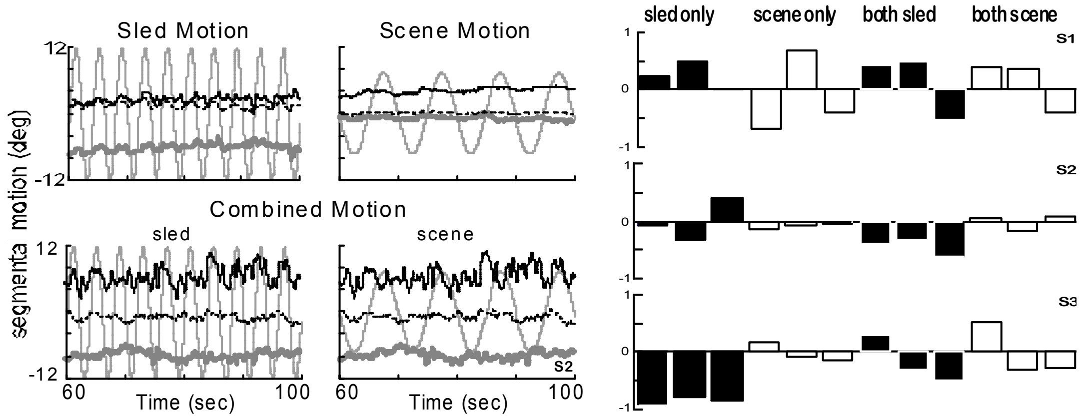


Figure 2. *(Left) Relative excursions of the head re trunk (bold grey line), trunk re shank (black line), and shank re sled (dotted line) in a young adult subject plotted over a 40 sec period of the trial for each condition The same data in the combined condition is overlies sled and scene (thin grey lines) motion. (Right) Overall weighting of the input variables (sled = filled bars; scene = open bars) derived from the PCA for 3 young adults. Each bar of each condition represents a subsequent non-overlapping 40 sec time period. The direction of each bar indicates the relative phase between the response and the input signal.*

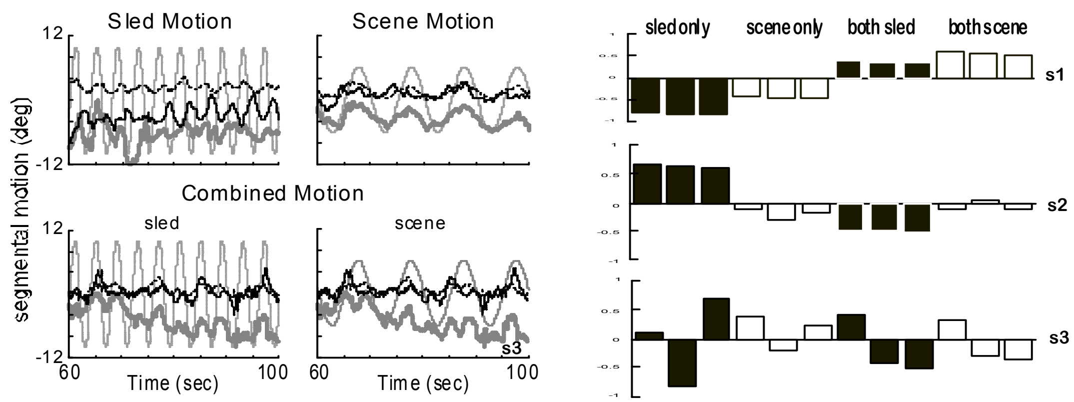


Figure 3. *(Left) Relative excursions of the head re trunk (bold grey line), trunk re shank (black line), and shank re sled (dotted line) in a 67 year old subject plotted over a 40 sec period of the trial for each condition The same data in the combined condition is overlies sled and scene (thin grey lines) motion. (Right) Overall weighting of the input variables (sled = filled bars; scene = open bars) derived from the PCA for 3 elderly adults. Each bar of each condition represents a subsequent non-overlapping 40 sec time period. The direction of each bar indicates the relative phase between the response and the input signal.*

3.2 Responses in Elderly Adults

Maintaining balance with either sled or scene motion was a challenge to the elderly subjects, and some subjects were unable to complete the trials when the two stimuli were combined. When only the sled moved, these subjects tended to exhibit a response of the trunk that was compensatory to the motion of the sled and a response of the shank that matched the direction of the sled (Fig. 3) from which we infer that they were bending at the hip. When only the scene moved, the elderly subjects were more affected by motion of the visual scene than were the younger adults. This was demonstrated by large responses of the head with respect to the trunk in the direction of scene motion (Fig. 3). This sensitivity to visual motion carried over into the combined protocol where the elderly subjects exhibited clear motion of all body segments in the direction of the visual scene motion. Greater equivalency of the two input signals emerged in the PCA results (see S1 and S3 in Fig. 3) in both the trials with a single input and in the trials with combined inputs.

3.3 Tracking a visual target

Differential weighting of sensory inputs suggests that an individual's perception of the stimulus array might influence their response. To further explore how visual attention modifies balance reactions, we asked both young and elderly subjects to perform a visual tracking task while the full field of view was translated fore-aft. When asked to focus on either a stationary or moving target, no differences were observed between the young and older adults. In both groups, button press latencies for identifying the correct number calculation were shorter during the non-computation tasks than during the computation tasks indicating that the task did convey an attentional load in both groups.

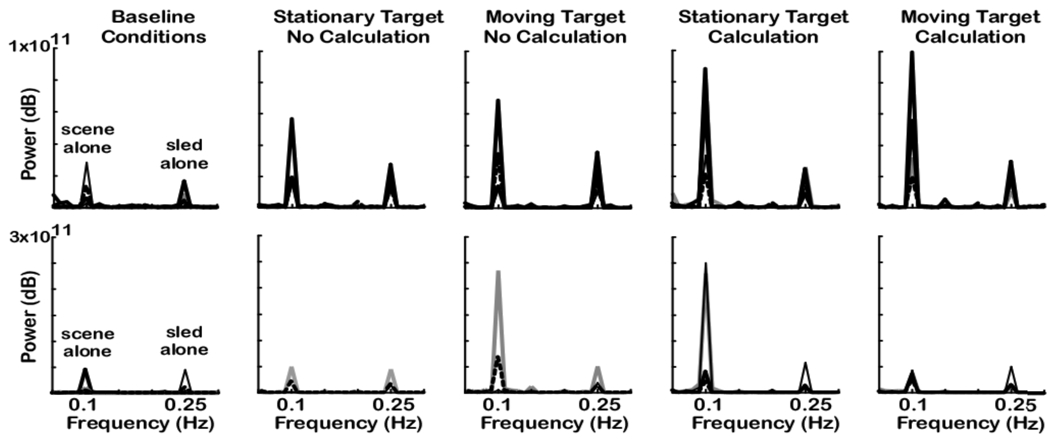


Figure 4. *Power derived from fast Fourier transforms performed on the COM response of the head for five young (top row) and four older (bottom row) adults with each visual task (each line is the response of one subject). Note the greater magnitude of the power scale for the older adults.*

When we examined the COM response at each body segment, we found that all subjects responded predominantly at the input frequencies of the sled (0.25 Hz) and the visual scene (0.1 Hz), but power of the response was much greater in the elderly subjects overall (Fig. 4). For both groups, power of the head increased only at the frequency of the visual scene when combined with the focal task. In the young adults, power at 0.1 Hz increased for the head when the visual target was moving, and increased even more when calculation was required whether the target was stationary or moving. For both groups, power at the head was significantly greater during the stationary calculation task than during the stationary no-calculation task ($p < 0.01$), although of the four elderly subjects, only two exhibited any real differences. The magnitude of head excursion also increased with a stationary calculation task when compared to the stationary no-calculation task ($p < 0.001$). Older adults exhibited greater head angular displacements than the young adults during the moving target tasks, but similar COM displacements suggest that they were using different kinematic strategies for head stabilization.

4. CONCLUSIONS

The results presented here argue that postural responses to motion of the visual field are strongly potentiated by the presence of both physical motion and an additional visual task. Either support surface or visual field motion alone produced marginal responses in most subjects, but the response to visual stimulation was dramatically enhanced when these inputs were combined, perhaps because the visual inputs were incongruent with those of the physical motion. Also, adding the calculation task to the combination of visual scene and support surface motion reduced the typically stabilizing effect of a stationary visual target (Strupp et al, 2003).

With single inputs, some subjects consistently selected a single segmental strategy. With multiple inputs, most subjects produced behaviours that fluctuated between the two stimuli. A non-additive effect occurred in the energy of the response with combined inputs in that the magnitude of the response kinematics greatly increased. Segmental kinematics in the older adults were more affected by visual scene motion than the young adults, implying that instability observed in elderly adults may be due to an inability to minimize their responses to inappropriate or irrelevant inputs. The dependence of the elderly subjects on stabilizing their posture by flexing at the hip also suggests that their more rigid and unstable trunk (Allum et al, 2002; Wu, 1998; Keshner, 2004) would be less capable of minimizing disturbances to the position of the head in space.

Segmental responses to visual field motion were also strongly potentiated when the visual field motion was combined with physical motion suggesting that there are greater consequences of engaging in a cognitive task when there are multimodal demands on posture (Jamet et al, 2004). During functional activities, postural control is a continuous process and an integral component of any activity we engage in. Our current results demonstrate that individual perception of the sensory structure was a significant component of the postural response. Previous results from our laboratory (Keshner and Kenyon, 2000; Keshner et al, 2004) have revealed time dependent properties in the postural control system. We suggest that a constant recalculation of the stimulus flow is required to appropriately tune the response to the environmental context. In elderly adults, the inability to actively suppress continuous regulation of the visual environment could result in a miscalculation between the required postural responses and relevant environmental demands.

These results have significant implications for the continued measurement of postural activity and the important role of the virtual environment in the research and rehabilitation environment. The adaptive nature of the human nervous system makes it imperative that we test and train individuals in conditions as close as possible to those they will encounter during their daily activities. Our system allows us to explore more complex behaviours that are necessary for rehabilitation (Keshner and Kenyon, 2004). We believe that application of VE technology to dynamic postural research is both a necessary and valid approach for exploring underlying control mechanisms and questions that are relevant to rehabilitation.

Acknowledgements: This work was supported by National Institute of Health grants DC01125 and DC05235 from the NIDCD and AG16359 from the NIA. We gratefully acknowledge VRCO (http://www.vrco.com) for supplying CAVE library software and the Ascension Tracker Daemon.

5. REFERENCES

J H Allum, M G Carpenter, F Honegger, A L Adkin and B R Bloem (2002), Age-dependent variations in the directional sensitivity of balance corrections and compensatory arm movements in man, *J. Physiol.*, **542**, pp. 643-663.

J Buchanan and F Horak (1999), Emergence of postural patterns as a function of vision and translation frequency, *J. Neurophysiol.,* **81**, pp. 2325-2339.

C Cruz-Neira, D J Sandin, T A DeFanti, R V Kenyon and J C Hart (1992), The CAVE automatic virtual environment, *Communications*, **38**, pp. 64-72.

J Dichgans, R Held, L R Young and T Brandt (1972), Moving visual scenes influence the apparent direction of gravity, *Science*, **178**, pp. 1217-1219.

L J Hettinger, K S Berbaum, R S Kennedy, W P Dunlap and M D Nolan (1990), Vection and simulator sickness, *Military Psychology*, **2**, pp. 171-181.

M Jamet, D Deviterne, G C Gauchard , G Vancon and P P Perrin (2004), Higher visual dependency increases balance control perturbation during cognitive task fulfillment in elderly people, *Neurosci. Lett.,* **359**, pp. 61-64.

E A Keshner (2004), Head-trunk coordination in elderly subjects during linear anterior-posterior translations, *Exp Brain Res.*, DOI: http://dx.doi.org/10.1007/s00221-004-1893-2.

E A Keshner, J H J Allum and C R Pfaltz (1987), Postural coactivation and adaptation in the sway stabilizing responses of normals and patients with bilateral peripheral vestibular deficit, *Exp. Brain Res.,* **69**, pp. 77-92.

E A Keshner and R V Kenyon (2000), The influence of an immersive virtual environment on the segmental organization of postural stabilizing responses, *J. Vestib. Res.*, **10**, pp. 207-219.

E A Keshner and R V Kenyon (2004), Using immersive technology for postural research and rehabilitation, *Assistive Technol.*, (in press).

E A Keshner, R V Kenyon and J Langston (2004), Postural responses exhibit intra-modal dependencies with discordant visual and support surface motion, *J. Vestib. Res.*, **14**, (in press).

L Nashner and A Berthoz (1978), Visual contribution to rapid motor responses during postural control, *Brain Res.,* **150**, pp. 403-407.

R J Peterka, F O Black and M B Schoenhoff (1990), Age-related changes in human vestibulo-ocular reflexes: Sinusoidal rotation and caloric tests, *J. Vestib. Res.,* **1**, pp. 49-59.

M Strupp, S Glasauer, K Jahn, E Schneider, S Krafczyk and T Brandt (2003), Eye movements and balance. *Ann. N. Y. Acad. Sci.,* **1004**, pp. 352-358.

G Wu (1998), Age-related differences in body segmental movement during perturbed stance in humans, *Clin. Biomech*., **13**, 300-307.

Virtual reality in the rehabilitation of the upper limb after stroke: the user's perspective

J H Crosbie[1], S M McDonough[2], S Lennon[3], L Pokluda[4] and M D J McNeill[5]

[1,2,3]Rehabilitation Sciences Research Group, University of Ulster,
Shore Road, Newtownabbey, Co. Antrim, BT37 0QB, NORTHERN IRELAND

[4,5]Faculty of Engineering, University of Ulster,
Cromore Road, Coleraine, Co. Londonderry, BT52 1SA, NORTHERN IRELAND

j.crosbie@ulster.ac.uk, *s.mcdonough@ulster.ac.uk*, *s.lennon@ulster.ac.uk*,
xpokluda@fi.mini.cz, *mdj.mcneill@ulster.ac.uk*

www.ulster.ac.uk

ABSTRACT

Virtual reality provides a three-dimensional computer representation of a real world or imaginary space through which a person can navigate and interact with objects to carry out specific tasks. One novel application of VR technology is in rehabilitation following stroke, particularly of the upper limb. Our research group has built a system for use in this field, which gives the user the ability to interact with objects by touching, grasping and moving their upper limb. A range of user perspectives has been tested with healthy individuals and with people following stroke.

1. INTRODUCTION

Virtual reality (VR) provides a three-dimensional computer representation of a real world or imaginary space through which a person can navigate and interact with objects to carry out specific tasks. It can either be *immersive*, where the user feels physically present in the virtual environment (VE), typically using a head-mounted display (HMD) or *non-immersive*, where a handheld interface allows interaction with objects on a computer screen. In its immersive form, visual, auditory and tactile sensory aspects of the VE can be delivered to the individual through visual display units and speakers within the HMD unit (Rose et al, 1996). A tracking system links head movements so that the image in the VE is updated in synchrony with head movements, giving an impression of looking around from within the VE. In addition to this, a representation of the user's upper limb can be generated within the VE using sensors on the arm and a dataglove on the hand (McDonough et al, 2003) so that the user can interact with objects as though they were real (Kozak et al 1993).

One novel application of VR technology is in rehabilitation following stroke, particularly of the upper limb. Recovery of upper limb function is a major problem, with 30 – 66 % of stroke survivors no longer being able to use the affected arm (van der Lee et al 1999). This can be explained in part by the site of injury in the cortex (Kandel et al, 1991), which can cause limb paresis that limits active practice with the arm in the real world (Chae et al, 1998). Other factors are low levels of interaction between the patient and the environment (Tinson, 1989; Mackey et al, 1996), ineffective therapy techniques (Kraft et al 1992), and the very small percentage of time actually spent practicing tasks (Tinson, 1989). The use of VE technology has been advocated for this problem as it can be manipulated to avoid the physical constraints that would prevent a patient practicing in the real world. The ability to practice is very important as it has been shown in the rehabilitation literature that early (Wade et al, 1985; Chae et al, 1998; Johansson, 2000) intensive (Langhorne et al, 1996; Kwakkel et al, 1997; Kwakkel et al, 1999) practice of active functional tasks (Dean and Mackey 1992, Smith et al, 1998; Smith et al, 1999) in an enriched environment (Bennett 1976; Rosenzweig, 1980, Ohlsson & Johansson, 1995; Grabowski et al. 1995) leads to more positive outcomes for upper limb rehabilitation, by modifying neural reorganisation of the cerebral cortex (Thompson et al, 1996; Traversa et al, 1997).

The specific advantages of both non-immersive and immersive VE technology in people following stroke are: the ability to view and correct the movement of the affected arm; to explore, interact and make errors, which provides the facility for motor learning. Tasks can be chosen to enable success that would not be possible in the real world; and their difficulty can be increased, over time, to mirror real world tasks. Immersive VE technology has the additional advantage of increasing the feeling of being 'present' in the VE and this greater degree of 'realism' may be an advantage in rehabilitation of people following stroke. Indeed effectiveness of a VE has often been linked to the sense of presence reported by the user, where presence is defined as the subjective experience of being in one place or environment, even when one is physically situated in another (Witmer and Singer, 1998).

Although there is much potential for the use of immersive VE in clinical applications, there are problems that could limit their ultimate usability. Some users have experienced side effects during and after exposure to an immersive VE. There have been some reports of users experiencing transient reduced binocular vision after wearing head-mounted displays (Mon-Williams et al, 1993). Other symptoms include disorientation and balance disturbances, and nausea. Many of these effects can be attributed to delays between the sampling of head and limb positions and the presentation of an appropriate image through the HMD. This incongruity between visual cues through the HMD and vestibular motion cues can produce sickness symptoms, similar to motion sickness, including disorientation, sweating, nausea and headache. It should be noted that there is considerable variation in the extent to which people in a VE have been found to suffer from these problems (Kolanski, 1995). Susceptibility to side effects can be affected by age, ethnicity, experience, gender and physical fitness, as well as the characteristics of the display, the virtual environment and the tasks (Lewis and Griffin, 1997).

The majority of published work on the use of VR technology for rehabilitation of motor deficits following stroke (n=11) in the upper limb (Burdea et al, 1997, Holden et al, 1999, Myers and Biering, 2000, Jack et al, 2001, Piron et al, 2001, Broeren et al, 2002, Deutsch et al, 2002; Holden & Dyer, 2002; Merians et al, 2002), and lower limb (Deutsch et al, 2002; Brown et al, 2002) have used non-immersive technology. No specific side effects have been noted in relation to the use of non-immersive VR with people after a stroke. Therefore the introduction of patients to immersive virtual reality environments, for assessment, therapy or rehabilitation, needed to be carefully tested in order to identify any safety or ethical issues.

Further questions that needed to be addressed in exploratory work of this system were: the specification of the tasks for the system from the users perspective; and the amount of effort and degree of immersion experienced using immersive VR during tasks in people following stroke.

Therefore the aim of this project was:

a) To ascertain the views of potential users of a virtual reality rehabilitation (VRR) system with respect to the type of task to be practised.
b) To establish the specification of these tasks to encourage arm and hand movement in people following stroke
c) To assess the interaction of the user, in both the healthy and stroke populations, in terms of their experience of presence in the virtual environment and their perceived exertion.
d) To investigate the rate of self reported side effects from use of the VRR system in both healthy and stroke users.

2. SECTIONS

2.1 Equipment: UUJ Virtual Reality Rehabilitation System

Our research group has built a system for use in stroke rehabilitation, which gives the user the ability to move around a world composed of familiar objects and to interact with these objects by touching, grasping and moving their upper limb. The user wears a head-mounted display (HMD) giving visual and audio cues and a data glove that facilitates manual interaction in the virtual world (see Figure 1).

A magnetic sensor system provides real-time 6-degrees of freedom (position and orientation) tracking of up to four points on the user's body. A VE has been created with a series of reaching and grasping tasks (based on the views from people following stroke, see Section 2.2); the patient sees a number of easily recognisable objects and also stylised representations of their arm and hand in the VE which replicates the movement of their upper limbs in the real world. The sensors attached to the upper limb are positioned carefully in order for the virtual arm to be correctly articulated. Three of the sensors are attached to the major upper limb joints (shoulder, elbow and wrist) and the fourth is attached to the HMD to facilitate the sense of immersion in the VE – as the patient moves their head their view of the VE changes accordingly.

Figure 1. *User wearing a head-mounted display (HMD) giving visual and a data glove that facilitates manual interaction in the virtual world.*

2.2 Identification of Tasks: Focus group interviews

To date three focus group interview sessions have been conducted by the research team. Participants were recruited from two local stroke clubs. The composition of the group members can be seen in Table 1 below, which details the individual's age, sex and relevant stroke characteristics.

Table 1. *Focus group participants.*

Group	Number	Sex (M=male, F=female)	Age (years)	Time since stroke (years)	Right/Left (R/L) stroke	Upper limb movement ✓=yes ✗=no
A	1	M	72	12	L	✓
	2	M	78	2	R	✓
	3	M	52	15	L	✓
	4	M	65	2	R	✓
	5	M	61	1	L	✗
	6	F	75	15	L	✗
B	7	F	78	7	R	✓
	8	F	62	9	R	✓
	9	M	63	6	L	✓
	10	M	73	3	L	✓
	11	M	62	12	R	✓
C	12	M	72	5	L	✓
	13	M	81	9 months	R	✓
	14	M	60	10	L	✓
	15	M	75	15	L	✓
	16	M	65	2	R	✗
	17	M	67	11	R	✓

Following the introductions and opening explanations of the project, participants were guided to discuss and give responses to the following questions:

1. What sort of activities can you do with your stroke arm?
2. What would you like to be able to do with your stroke arm?
3. What type of exercises / tasks do you think people with a stroke should be practicing with their arm?
4. How many of you had any form of therapy after your stroke?
5. Can you remember how much time was directed to getting movement back in your arm?
6. Do you think you would like to use a virtual computer system to help you practice arm movements?

The interviews were tape recorded and reviewed by the researchers. The main themes in relation to the above questions are summarised in Table 2.

Table2. *Summary of participant responses during focus group sessions.*

Question	Response summary
1 What sort of activities can you do with your stroke arm?	Reaching tasks Everyday actions
2 What would you like to be able to do with your stroke arm?	To hold objects To be able to use arm in getting dressed
3 What type of exercises / tasks do you think people with a stroke should be practicing with their arm?	Based on everyday actions Working above shoulder level A minority noted fine finger movements
4. How many of you had any form of therapy after your stroke?	The majority of participants had some form of therapy following their stroke, delivered via a mixture of in and out patient settings and ranging between 6 weeks and 5 months duration.
5. Can you remember how much time was directed to getting movement back in your arm?	The majority of participants reported the focus of therapy to be on the lower limb and walking practice. A minority remembered practising upper limb exercises.
6. Do you think you would like to use a virtual computer system to help you practice arm movements?	15 of the 17 participants stated they would like to try using a virtual computer system. A few were concerned as to their lack of knowledge about computers.

From the information gleaned from the focus group interviews the design of the UUJ VRR system, as described above, has attempted to incorporate the features that potential users would like to see in a rehabilitation system for the upper limb. In summary these are:

- Everyday activities
- Reach and retrieve actions
- Working above shoulder level
- A few wanted to focus on fine finger movements

A limitation of the above analysis is that at the time of interview most of the participants were between 2 and 15 years after having their stroke. Therefore the individual's recollection of how much therapy they received and how much of that was directed towards the rehabilitation of their upper limb specifically may not be reliable. Also, due to the chronic nature of most of these participants stroke, the type of exercises they would want to practice is likely to influenced and may be more limited in comparison to a group of people in the acute phase after stroke. It is planned to continue this analysis with focus group interviews with people in the acute phase and investigate of they have different ideas or perceptions. The VRR system may then need to be modified to incorporate alternative tasks.

2.3 Specification of Tasks

Functional tasks have been designed to incorporate a range of levels of difficulty for reach, grasp, release and manipulative components. The tasks engage both individual joint movements and the use of the whole arm and some require fine upper limb and hand motor control. They have been carefully designed so that complex tasks (such as grasping an object and moving it from one position to another) can be divided into a number of smaller sub-tasks. A wrist extension task has been designed to focus particularly on the movement at this joint (see Figure 2). It is likely to be functionally useful to promote wrist extensor activity as it impacts on grip strength and the dynamic positioning of the hand for grasping objects.

(i) Reach and Retrieve Task. This task has 7 levels of difficulty and will be tailored for each individual patient. The components of the task are 1) reach for a near green square object on a table in the VE, 2) then reach for and grasp a red cup in the centre of the table and 3) finally place the red cup on top of a yellow square object on the table. The tasks can be chosen depending on the patient's motor dysfunction, for example if the patient has difficulty reaching forward with the arm, the target for reaching 1) is increased in size and the distance to reach reduced. As the patient's reaching function improves, this task can be made progressively more difficult by reducing the size of the objects and increasing their distance away from the patient. The level of difficulty of the task can also be increased by increasing the number of objects that the patient has to touch, and by randomly ordering the placement of objects on the screen. Actual grasping of the cup 2) can either be produced on screen by simply touching the cup with any part of the hand (grasping is not required) to higher-level function where the hand must grasp around the cup in order to hold it.

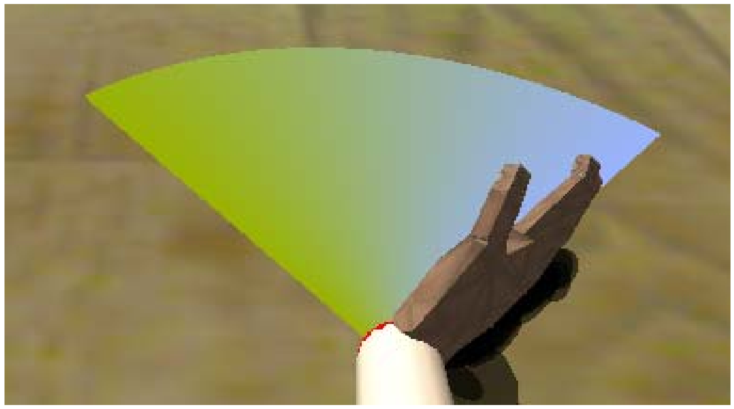


Figure 2. *View of the wrist extension task.*

(ii) Wrist extension. As noted above, the VE system can also be used to facilitate wrist extension practice. On screen the patient views a red task bar, each time they produce a flicker of wrist extensor muscle activity the colour of the task bar changes incrementally to green, if they carry out 10 successful repetitions the task bar changes to green completely, this can be repeated many times and alteration of the number of repetitions to complete can be reset depending on the users ability. The settings for this task can be changed so that only a very minor amount of movement starts the task, or the patient must produce moderate or full range of wrist extension in order to completely move the full red to the full green task bar.

Each sub task is accompanied by a set of audible instructions delivered via the earphones on the HMD and a pre-recorded demonstration for the user to view prior to attempting the task. The time taken to complete each task can be monitored, as can the patient's movements as they are engaged in achieving their goals. A specially designed graphical user interface (GUI) has been incorporated to allow straightforward initialisation of the system, configuration of individual tasks and scripting of whole sessions (see Figure 3), the goal being that this complex system should be usable by physiotherapists and other non-computing personnel.

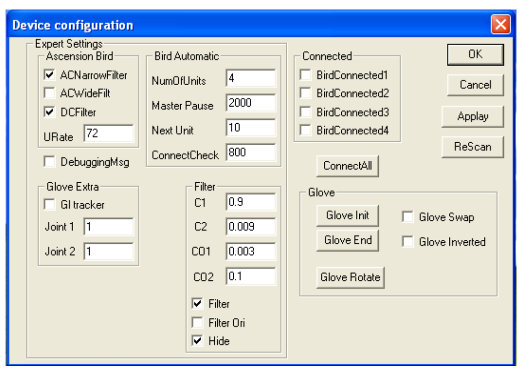


Figure 3. *System Initialisation and Task Configuration Dialogues*

2.4 User experience of the UUJ VRR system: Subjects

For this part of the study both healthy volunteers and people after stroke were recruited in order to assess their experience of using the VRR system. *Inclusion criteria* for those following stroke was: first stroke with a motor impairment of the upper limb as a primary deficit, muscle strength greater than 2/5 on the Medical Research Council (MRC) scale; stable medical condition; ability to communicate; good cognitive ability as indicated by a score greater than 15 on the Mini Mental State Examination (MMSE, Folstein et al 1975), Motricity Index score (Wade, 1992) greater than 26 indicating either activity of the shoulder or elbow and beginnings of prehension in the hand. *Exclusion criteria:* Patients with significant dysphasia; severe symptomatic arm pain; poor muscle strength (as determined by a score of less than 2/5 on the MRC scale of shoulder girdle muscle strength), an unstable medical condition, poor cognitive status (as defined by the MMSE, Folstein et al 1975) and severe visual-spatial neglect (The Line Cancellation Test, Albert 1973). Informed written or witnessed verbal consent was obtained from all recruited patients and healthy adults.

Each person undertook a VR session in which the HMD, dataglove and motions sensors were applied and they worked through a short series of exercises for the upper limb, which included reach to target tasks, a wrist extension task and a reach and retrieve of an object task. Features that have been designed into the system were evaluated, i.e. the audible instructions and pre-recorded demonstrations, the options to add or remove objects and to adjust distances and the height of objects in relation to the user. (see Figure 4 below).

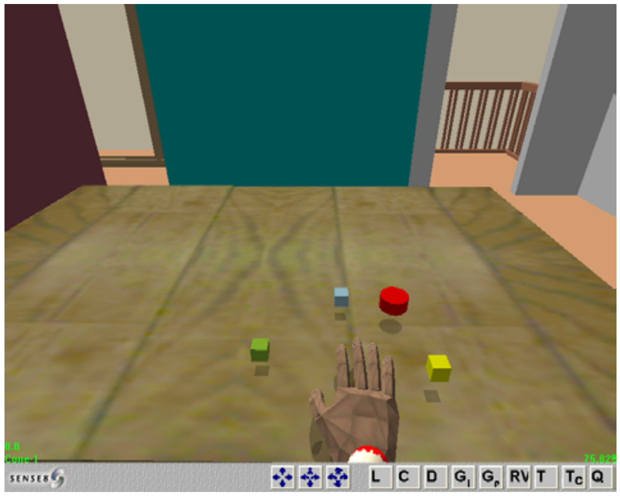


Figure 4. *View of one set up of the reach to target tasks.*

The subjects were divided into healthy and post stroke groups, with 10 and 5 individuals in each group respectively. The composition of the groups in terms of age, sex and details of stroke are displayed in Table 3. The mean age of the healthy group was 42 years and of the stroke group 62 years. Stroke users may have right or left sided weakness and thus the healthy group used a mixture of their right or left hands to operate the VR system to reflect this. The time since stroke of this group ranged from 6 days to 25 years, the mean being 10 years. Again this group comprised more people with chronic stroke as compared to those in the acute phase. The Immersive Tendencies Questionnaire was administered prior to the subject participating in the VR session (see Section 2.4.1). The Task Specific Feedback Questionnaire (see Section 2.4.1) and the Borg Scale of Perceived Exertion (see Section 2.4.2) was administered immediately after the VR session.

2.4.1 Presence. The Immersive Tendencies Questionnaire (ITQ) was developed by Witmer and Singer (1998) and was used in this study to determine differences in the tendencies of individuals to experience presence.

The questionnaire asks the participant to answer 29 questions about issues such as how involved they become in doing tasks, when watching television or movies and also if they are easily distracted from tasks. An example is displayed below:

Do you ever become so involved in doing something that you lose all track of time?

Never Always

The scale is similar to a visual analogue scale for rating pain perception. From their answers to each question an impression is gained about how susceptible this person might be to being immersed in the VE. The person is scored as having a positive or negative immersive tendency. There is a positive relationship between presence and task performance. Thus a person with a positive immersive tendency might be more likely to be successful in the performance of virtual tasks. It can be seen from the summary of results in Table 4 that in the healthy adults 50% scored positively on the ITQ and 50% scored negatively, indicating that there was an equal balance of those more likely to become immersed as those less likely to become immersed in the VE. A similar pattern was seen on the people with stroke, although data for the outcome has only been measured for 3 subjects to date.

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

Table 3. *VRR testing participants.*

Healthy group			
Subject number		Age (years)	Right / Left handed use of the VR system
1		70	Right
2		48	Left
3		39	Right
4		44	Right
5		42	Left
6		48	Right
7		37	Left
8		23	Right
9		44	Right
10		28	Left
Post stroke group			
Subject number	Age (years)	Right / left stroke	Time since stroke
1	65	Left	3 years
2	83	Right	6 days
3	78	Right	2 years
4	62	Right	9 years
5	25	Right	25 years

2.4.2 User Experience. The Task Specific Feedback Questionnaire (TSFQ, Kizony et al, 2003) is a modified version of the Witmer and Singer (1998) Presence Questionnaire and queries the users perceived difficulty of the tasks carried out in the VE. This was used to evaluate the experience of using the system. The instrument comprises 6 questions responded to on a scale of 1 – 5, giving a total score of between 1 and 30. For example:

1. Did you experience any feeling of enjoyment?

1 2 3 4 5

1 = strongly agree, 2 = agree, 3 = unsure, 4 = disagree and 5 = strongly disagree

Lower scores indicate that the user had a more favourable experience using the virtual system.

Users of visual simulators may experience simulator sickness symptoms. These can include discomfort, fatigue, headache, nausea and dizziness (Kennedy et al, 1993). Question 7 on the TSFQ asks the participant to indicate if they experienced any discomfort or not. The researcher further asked participants to self-report the nature of the discomfort or the experience of simulator symptoms following the use of our VRR system; this is reported in Section 2.4.3. It can be seen from Table 4 that similar scores on the TSFQ were obtained for individuals in both the healthy and stroke groups, with a mean score of 10.8 and 14.8 for each group respectively. The scores are at the lower end of the 0 – 30-point scale, which as reported above is likely to indicate that the person had a more favourable experience when immersed in the VE. There was one erroneous result from subject 5 in the stroke group with a TSFQ score of 28.

2.4.3 Exertion and Side Effects. The Borg Scale of Perceived Exertion (Borg, 1982) was used to assess the individual's perception of any physical exertion they may have experienced whilst exercising their arm in the VE. Perceived exertion is the overall effort or distress of the body during exercise. The number 0 represents no perceived exertion or discomfort and 10 represents the greatest amount of perceived exertion that the person has ever experienced. Table 4 includes the Borg ratings for each group. It can be seen that in this case there is a difference between the healthy participants and the people with stroke. The subjects in the healthy group rated their perceived level of exertion as being weak, with scores of between 0 and 4. This equates to a range of nothing at all to moderate level of perceived exertion. However, the people in the stroke group rated their experience between 3 and 7 on the Borg scale. This equates to a range of moderate to very strong perceptions of exertion. This is probably to be expected as for individuals with no upper limb motor deficits the VR tasks are quite easy to execute; whereas, for the person with stroke the tasks are likely to require an increase in motor activity and thus effort to complete them.

Table 4 *Results summary of the Immersive Tendencies Questionnaire (ITQ), Task Specific Feedback Questionnaire (TSFQ), and Borg Scale of Perceived Exertion for healthy and stroke groups.*

Healthy group				
Subject number	ITQ (positive / negative)	Borg rating	TSFQ score	Reported side effects (yes / no)
1	Negative	2 - weak	13	No
2	Negative	2 - weak	12	Yes
3	Negative	1 – very weak	9	Yes
4	Positive	1 – very weak	8	Yes
5	Positive	0 - nothing	12	No
6	Negative	0.5 extremely weak	9	No
7	Positive	2 - weak	6	Yes
8	Positive	0.5 extremely weak	11	No
9	Negative	4 – weak/moderate	17	Yes
10	Positive	3 - moderate	11	No
Post stroke group				
Subject number	ITQ (positive / negative)	Borg rating	TSFQ score	Reported side effects (yes / no)
1	Positive	7 – very strong	12	Yes
2	Negative	5 – strong	9	No
3	Positive	3 – moderate	10	No
4	No data	5 – strong	15	No
5	No data	5 - strong	28	No

With respect to side effects experienced having been immersed in a VE, of the healthy group 5 out of 10 experienced transient symptoms, which included headache, dizziness, discomfort and nausea. Only one stroke user experienced symptoms of dizziness, which again was transient.

3. CONCLUSIONS

The VRR system produced by the research team has been tested with respect to the user experience in a small group of healthy individuals and people with stroke. It has been shown that amongst both groups there was an equal balance of subjects more or less likely to become immersed in a VE as tested by the Immersive Tendencies Questionnaire. Participants in both groups indicated a generally favourable experience using the VRR system, as indicated by the lower scores on the Task Specific Feedback Questionnaire. Similarly some healthy users and one stroke user reported transient side effects following their interaction with the system. The main difference between the two groups tested was their level of perceived exertion, as measured by the Borg Scale. The healthy group rated the experience of using the VRR rehabilitation system as none to a weak level of exertion. Whilst the users with stroke perceived the level of exertion demanded to be at a higher level of moderate to strong.

Acknowledgements: Northern Ireland Chest, Heart and Stroke Association, 6th Floor, 22 Great Victoria Street, Belfast BT2 7LX (Application number: 2002004).

4. REFERENCES

M L Albert (1973,) A simple test of visual neglect, *Neurology*, **23**, pp. 658-664.

E L Bennett EL (1976), *Neural Mechanisms of Learning and Memory*. MIT Press Ltd: Cambridge, MA.

G V Borg (1982) Psychophysical bases of perceived exertion, *Med Sci in Sport and Exe*, 14, pp. 377-381.

J Broeren, A Bjorkdahl, R Pascher and M Rydmark (2002), Virtual reality and haptics as an assessment device in the post acute phase after stroke, *Cyber Psych Behav*, **5**, pp. 207 – 211.

D A Brown, D L Jaffe and E L Buckley (2002), Use of virtual objects to improve gait velocity in individuals with post stroke hemiplegia, *Neurology Report*, **26**, pp. 105.

G Burdea, S Deshpande, N Langrana, D Gomez and B Liu (1997), A virtual reality-based system for hand diagnosis and rehabilitation, *Presence – Teleoperators and Virtual Environments*, **6**, pp. 229 – 248.

J Chae, F Berthoux and T Bohine (1998), Neuromuscular stimulation for upper extremity motor and functional recovery in acute hemiplegia, *Stroke,* **29**: pp. 975-979.

C Dean and F Mackey (1992), Motor assessment scale scores as a measure of rehabilitation movement control following stroke, *Physio*, **69**, pp. 238-240.

J E Deutsch, A S Merians, S Adamovich, G Burdea and H Poizner (2002), Haptics and virtual reality used to increase strength and improve function in individuals post-stroke, *Neur Report***,** **26**, 2, pp. 79-86.

M Grabowski, J C Sørensen, B Mattson, J Zimmer and B B Johansson (1995), Influence of an enriched environment and cortical grafting on functional outcome in brain infarcts in adult rats, *Experimental Neurology*, **133**, pp. 1-7.

M F Folstein, S E Folstein and P R McHugh (1975), Mini Mental state. A practical method for grading the cognitive state of patients for the clinician, *Journal of Psychiatric* Research, **12**, pp. 189-198.

M A Grealy, D A Johnson and S K Rushton (1999), Improving cognitive function after brain injury: The use of exercise and virtual reality, *Arch Phys Med Rehab,* **80**,pp. 661-667.

M Holden and T Dyar (2002), Virtual environment training – a new tool for neurological rehabilitation, *Neur Report***,** **26**, 2, pp. 62-71.

M Holden, E Todorov, J Callahan and E Bizzi (1999), Virtual environment training improves motor performance in two patients with stroke: case report, *Neur Report*, **23**, pp. 57 – 67.

D Jack, R Boian, A S Merians et al (2001), Virtual reality-enhanced stroke rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* **9**, pp. 308 – 318.

B B Johansson (2000), Brain plasticity and stroke rehabilitation, *Stroke*, **31**, pp. 223-230.

E R Kandel, J H Schwartz and T M Jessell (1991), *Principles of neural science* (3rd edition), Prentice-Hall International Inc, London.

R S Kennedy, N E Lane, K S Berbaum and M G Lilienthal (1993), Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness, *Intern J Aviation Psycho*, **3**, 3, pp. 203-220.

R Kizony, L Raz, N Katz, H Weingarden and P L Weiss (2003), Using a video projected VR system for patients with spinal cord injury, *Proc 2nd International Workshop on Virtual Rehabilitation*, Piscataway, pp. 82-88.

E M Kolanski (1995), Simulator sickness in virtual environments. *Technology Report 1027:* US Army Research Institute for the Behavioural and Social Sciences Virginia.

J J Kozak, P A Hancock, E J Arthur and S T Chrysler (1993), Transfer of training from virtual reality. *Ergonomics*, **36**, pp. 777 – 784.

G H Kraft, S S Fitts and M C Hammond (1992), Techniques to improve function of the arm and hand in chronic hemiplegia, *Arch Phys Med Rehab,* **73,** 3, pp. 220-227.

G Kwakkel, R C Wagenaar, T W Koelman, G H Lankhorst and J C Koetsier (1997), Effects of intensity of rehabilitation after stroke: a research synthesis, *Stroke*, **28**, pp. 1550-6.

G Kwakkel, R C Wagenaar, J W R Twisk, G H Lankhorst and J C Koetsier (1999), Intensity of leg and arm training after primary middle-cerebral-artery stroke: a randomised trial, *Lancet (North American Edition)*. **354**, pp. 191-6.

P Langhorne, R Wagenaar and C Partridge (1996), Physiotherapy after stroke: more is better? *Phys Res Intern,*. **1,** 2, pp. 75-88.

C H Lewis and M J Griffin (1997), Human factors consideration in clinical applications of virtual reality. *Studies in Health Technology and Informatics,* **44**, pp. 35 – 56.

F Mackey, L Ada and R Heard (1996), Stroke rehabilitation: are highly structured units more conducive to physical activity than less structures units? *Arch Phys Med Rehab*, **77**, pp. 1066-1070.

S M McDonough, J H Crosbie, L Pokluda and M D J McNeill (2003, The use of virtual reality in the rehabilitation of the upper limb following hemiplegic stroke, *Proc 2nd International Workshop on Virtual Rehabilitation*, Piscataway, pp. 98.

A S Merians, D Jack and R Boian et al (2002), Virtual reality-augmented rehabilitation for patients following stroke, *Phys Ther* **82**, pp. 898 – 915.

M A Mon-Wiliams, J P Wann, S K Rushton and R Akerley (1993), Real problems with virtual worlds. *Ophthalmic Physiological Optics,* **13**, pp. 435 – 436.

R L Myers and T Bierig (2000), Virtual reality and left hemineglect: a technology for assessment and therapy, *Cyberpsychology and Behaviour*, **3**, pp. 465 – 468.

A-L Ohlsson and B B Johansson (1995), The environment influences functional outcome of cerebral infarction in rats, *Stroke*, **26**, pp. 644-649.

L Piron, F Cenni, P Tonin and M Dam (2001), Virtual reality as an assessment tool for arm motor deficits after brain lesions, *Stud Health Tech and Inform*, **81**, pp. 386 – 392.

F D Rose, D A Johnson, E A Attree, A G Leadbetter and T K Andrews (1996), Virtual reality in neurological rehabilitation, *Brit J of Ther and Rehab,* **3**, pp. 223 – 228.

M Rosenzweig (1980), Animal models for effects of brain lesions and for rehabilitation, In *Recovery of function: Theoretical considerations for brain injury*. University Park Press. Baltimore, MD.

G V Smith, R F Macko, K H C Silver and A P Goldberg (1998), Treadmill aerobic exercise improves quadriceps strength in patients with chronic hemiparsis following stroke: a preliminary report, *J of Neurologic Rehab*, **12**, 3, pp. 111-8.

G V Smith, K H C Silver, A P Goldberg and R F Macko (1999), "Task-oriented" exercise improves hamstring strength and spastic reflexes in chronic stroke patients, *Stroke,* **30**, pp. 2112-8.

M L Thompson, G W Thickbroom, BA Laing, S A Wilson and F L Mastaglia (1996), Transcranial magnetic stimulation studies of the corticomotor projection to the hand after sub-cortical stroke, *Movement Disorders*, **11**, pp. 25.1.

D J Tinson (1989), How stroke patients spend their days, *Int Disab Studies*, **11**, pp. 45-49.

R Traversa, P Cicinelli, A B Bassi, P M Rossini and G Bernardi (1997), Mapping of motor cortical reorganization after stroke, *Stroke*, **28**, pp. 110-117.

J H van der , R C Wagenaar and G J Lankhorst (1999), Forced use of the upper extremity in chronic stroke patients, *Stroke,*. **30**, pp. 2369-2375.

D T Wade (1992), *Measurement in neurological rehabilitation.* Oxford: Oxford University Press.

D T Wade, V A Wood and R Langton Hewer (1985), Recovery after stroke – the first three months, *J of Neuroy, Neurosurg and Psych*, **48**, pp. 7-13.

B G Witmer and M J Singer (1998), Measuring presence in virtual environments: A presence questionnaire. *Presence,* **7**, pp. 225 – 240.

Participants responses to a stroke training simulator

M Maxhall[1,2], A Backman[2], K Holmlund[2], L Hedman[3], B Sondell[1] and G Bucht[1]

[1]Department of Community Medicine and Rehabilitation, Geriatric Medicine, Umeå University, Umeå, SWEDEN

[2] VRlab, Umeå University, Umeå, SWEDEN

[3]Skill Acquisition Lab, Umeå University, Umeå, SWEDEN

[1,2]masmal@hpc2n.umu.se, [2]andersb@cs.umu.se, [2]Kenneth.Holmlund@hpc2n.umu.se, [3]leif.hedman@psy.umu.se, [1]bjorn.sondell@germed.umu.se, [1]gosta.bucht@germed.umu.se

[1]www.umu.se/medfak/institutioner/samh_reh_eng.html, [2]www.vrlab.umu.se, [3]www.psy.umu.se/forskning/Forskare/l_hedman

ABSTRACT

The primary goal of this research was to study a virtual environments (VE) possibility to influence empathy on caregiver personal. In the present explorative study, 9 subjects from Norrlands University Hospital (NUS) completed a test consistent of three everyday tasks, reading a newspaper, filling a glass of water and putting toothpaste on a toothbrush. The procedure was done twice first from a non-stroke perspective and secondly from a perspective of a patient with stroke handicaps. The VE looked like a normal apartment and could be experienced with or without different perceptual disorders of stroke. Data from interviews and observations was analyzed via methods inspired by Grounded Theory. Results from observations and interviews indicate that the simulator in spite of problems of usability were effective in influencing caregivers empathy.

1. INTRODUCTION

Stroke is one of the most common diseases. Costs related to stroke, is the most costly of all in healthcare budgets in the industrialized world. In Sweden about 35000 persons is affected every year, The yearly cost for society is estimated to 9.9 billion SKR (€ 1.1 billion) (Ghatnekar et al, 2004). The individual problems caused by stroke can vary greatly depending on the location and severity of the brain damage, and the degree of support from the surroundings. Symptoms may include; motor problems such as poor balance and limb control, problems with the field of vision, other problems related to emotional control and behaviour. These stroke related problems therefore one way or another directly influence the persons activities of daily living. A previously simple task such as going to the bathroom may now need assistance from others, or new strategies may have to be employed, initially often through the support from others. The affected individual may therefore suffer problems on several different levels. Due to the impact and character of the damage, stroke could secondarily create conflicts within the meeting with family, friends, home caring and living style. (The National Board of Health and Welfare, 2004).

In general there is a need to develop methods which may improve the situation for stroke affected individuals, their friends and families. In stroke rehabilitation efforts are made to meet the individual in his or her current total situation and work from there. Each individual may have her own background story and her own expectations and goals in life both before and after being affected by the stroke. There may be great individual differences regarding social support and social abilities and cognitive function.

Concomitant disease is another important issue regarding stroke and stroke rehabilitation, for some individuals there may be sufficient cardiovascular disease which may hinder parts of the stroke rehabilitation training. Visual impairment may make the use of motorized wheelchair impossible when the wheelchair is user driven, here the use of assistant guided motorised wheelchairs may provide sufficient transportation for the assistant but may not yield the satisfaction wanted for the patient who wants to drive as he has always

©2002 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

done. Understanding these issues may be a cue to better relating to the patients real problem which may be feeling sorrow for the loss of the ability to drive independently on his own. Here the use of tight rehabilitation teams with a multiprofessional approach is often a method used. Many different categories of professionals provide their support in a network around the patient. For this setup to work properly around the patient and patient family, there is a need for a tight well-functioning network of dedicated professionals, with a deep understanding for the patient related issues, in order to provide the best possible professional support.

Continuous professional education and development is crucial, but oftentimes there are only limited budgets today for such professional education. Sometimes relatives and caregivers can together take part of the rehabilitative training process. Especially relatives may find this part of the support very helpful. Many times there is certain levels of performance seen during training at the hospital or at home when trained personnel are present that may not necessarily be evident in the persons own home among relatives and vice versa. Both relatives, the patient and the caregivers can therefore learn from each other on how to address certain issues and how take new steps to move forward in the rehabilitative process. Here sometimes ways are also found on how to move forward in life in general. The importance of being able to understand each other and gaining a mutual feeling for each other on several levels, among these empathy, is important to reach as far as possible in improving life for all involved.

Improving empathy might thus be an important goal in order for relatives and care personnel to be better careers. Empathy has been described in several ways (Holm, 1995, 2001), most commonly related to an interaction between individuals. Empathy is characterized as an ability to place oneself mentally and emotionally in the world of another person, be sensitive to another's state of mind, current feelings and to communicate this understanding back to the other. The ability to be empathic depends on several factors, such as personality, health, attitudes and workload. According to Rogers (1975) unless you care about a person's wellbeing you can not accurately understand, value and sense the other person perceptions of the world. The term empathy was coined in 1909 and researchers has historically emphasized different aspects and character of empathy; cognition, affects, emotions, morality, individual similarity, understanding, altruistic behaviour and the relationship between empathizer and the object of empathy (Håkansson, 2003). However most researchers agree on that empathy is an active process consistent of all above-mentioned aspects (Holm, 1985, 2001). Empathy has been investigated and a positive outcome of training empathy among students and health care workers has been seen (LaMonica, 1976). If a simulator has the same effect, i.e. to improve persons empathic ability toward handicapped people is however not known.

Virtual Reality (VR) therapy, as a new therapeutic approach, that can be used to overcome some of the difficulties inherent in the traditional treatment of phobias. VR, like current imagine and in vivo modalities can generate stimuli that could be utilized in desensitization therapy. VR is being used to provide exposure and desensitization for a number of phobic conditions (Wiederhold et al, 2002. Roy et al, 2003. Garcia-Palacios et al, 2002. Wald et al, 2000, Vincelli et al, 2003)

In fear of flying it has been shown that a Virtual environment (VE) was more effective than an imagined environment. It is also suggested that physiological feedback may add to the efficacy of VR treatment. The physiological response of phobic participants, who were able to fly without medicine after VR treatment, showed a gradual trend toward the non-phobic's physiological responses as therapy sessions progressed (Wiederhold et al, 2002). Psychological interventions based on exposure therapy have proved to be effective, but given the particular nature of this disorder they bear important limitations.

Exposure therapy for fear of flying might be excessively costly in terms of time, money, and efforts. VR overcomes these difficulties as different significant environments might be created, where the patient can interact with what he or she fears while in a totally safe and protected environment-the therapist's consulting room (Banos et al, 2002).

Studies have also shown an effectiveness of low-budget VR exposure versus exposure in vivo in a between-group design in patients suffering from acrophobia. A study, for example, shows that VR exposure can be effective with relatively cheap hardware and software on stand-alone computers currently on the market (Emmelkamp et al. 2002).

Virtual reality training has many advantages over other clinical rehabilitation methods, and has the potential to develop a human performance training and testing environment (Lee et al, 2003).

Few studies have been made focusing on VR and empathy. Empathy has been studied as a cognitive factor and its relation with presence. It was concluded that empathy and creative experience, absorption and fantasy, play a distinctive role considering experiencing presence (Sas et al, 2002). Another study, not specifically investigating VR but the relationship between empathy and online interpersonal trust concludes that empathy and the accuracy of the empathic response between community members is an influential factor, but complicated and fragile (Feng et al, 2004).

It seems likely that VR could be an effective tool to influence and understand the problematic world of stroke patients and enhancing empathy.

2. AIMS AND METHODS

The main purpose with this study was to investigate if a virtual environment (VE), simulating stroke can influence caregiver's empathy for stroke patients and a stroke patient's daily experiences in their home environment.

In the present explorative study, 12 volunteers (2 physicians, 5 registered nurses, 3 licensed practical nurses, 1 nurse's aid and 1 occupational therapist) participated from the University Hospital of Northern Sweden (NUS). Nine of these (1 man and 8 women) completed the tests. Three nurses could not be tested due to technical problems. The tests took place during working hours, without reimbursement. The variation of age was between 22 and 52 years. Their experience of working with stroke was between 6 months and 16 years. Previous VE experience was non-existent.

The VE tests consisted of three everyday tasks, reading a newspaper, filling a glass of water and putting toothpaste on a toothbrush. The test procedure was repeated twice, initially from a non-stroke perspective and secondly from a simulated stroke-perspective with anomalies such as motion blur, unilateral neglect, and vestibular damage.

The VE looked like a typical Swedish apartment and could be experienced with or without different virtual perceptual disorders resembling a stroke, while navigating through it with a wheelchair as an interface.

3. TECHNICAL EQUIPMENT

The technical system was divided in two parts, hardware and software. The hardware system consisted of six parts, sensor system, wheelchair system, sound system, display system and two computers (one SGI Onyx 2 and one PC). The participant was during the trial equipped with an Ascension Motion Star tracker system, with nine motion sensors (hands, head and body). The system delivered sounds from objects in the VE, presented via a V8 head mounted display (HMD).

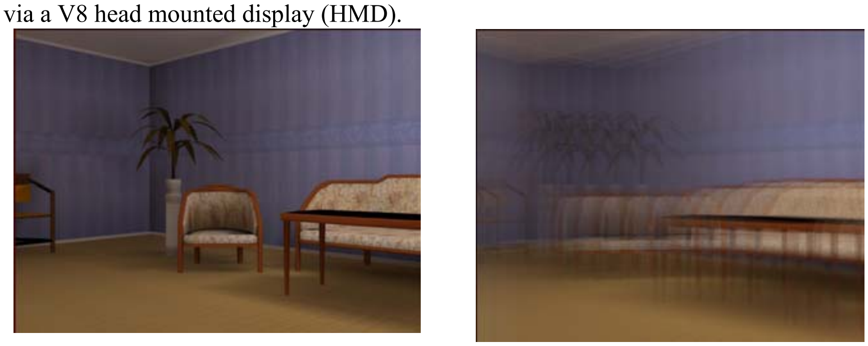


Figure 1. *Virtual apartment from a normal perspective (left image) and from a stroke perspective (right image).*

3.1 System Design

The participant is seated in an ordinary wheelchair which is mounted on a stand. This allows the wheels be elevated from the ground so that each wheel can freely rotate when the participant turns the wheels. Each wheel is equipped with step counters which measures the rotation. From this measurement the velocity of the wheels at any given time can be calculated, and hence the participant can use the wheelchair to navigate the VE.

A 3D model of an ordinary Swedish apartment consisting of a kitchen, a bathroom and a living room was used as the environmental model in the trials (Figure 1). This apartment also contains objects with interaction possibilities. The objects and their interaction possibilities were: glasses that could be filled with water from a tap; a newspaper that can be picked up; toothpaste that can be put on a toothbrush. To make the VE as natural as possible, the objects were also associated with sound events, such as streaming water etc. The participant was immersed in the VE environment using the HMD equipped with headphones.

On each hand the participant wear Pinch Gloves used for grasping objects in close proximity to each hand. There are four motion sensors used in the system: one on each hand of the participant, one mounted on the HMD and one attached to the wheelchair. The sensor on the wheelchair is used to calculate the participant's relative position to the wheelchair. The participant is free to navigate the apartment model using the physical wheelchair as the navigation device. The view and the interaction/navigation can be altered in the simulator to emulate the anomalies a stroke patient can suffer from. There are currently three stroke anomalies implemented in the simulator.

3.1.2 Anomalies. The anomalies are first described with the sensation experienced by real stroke patients, and later how they are implemented in the simulator.

3.1.3 Unilateral neglect. The sensation is of not being able to use the left side of the body (motor). It also implies an impairment of the left visual field (visual). Difficulties in finding objects to the left can occur, and also there may be difficulties in finding the proper way around in the apartment, when doors on the left might disappear. The implementation is focused on the navigation and the interaction aspect of the neglect. When this anomaly is activated, the left hand is switched of visually and for interaction. This means that only the right hand is available for interaction, mimicking the appearance of a stroke patient with neglect. Also, the camera in the VE which is controlled by the participants head is constrained to rotations to the right. This is also true for the wheelchair rotation which can only be navigated straight forward or to the right. When the participant is given the task of navigating a corridor and to move into a room located to the left, a 270° rotation to the right is the only way to solve the task, resulting in the same behaviour which can be observed in stroke patients suffering from neglect.

3.1.4 Motion blur. Motion blur is the sensation, described by patients, when objects and the surroundings perceived looking from one side to another. This anomaly is implemented using the accumulation buffer on the graphics computer. The amount of motion blur is parameterized and can be controlled during the setup of the test setup.

3.1.5 Vestibular damages. Vestibular damages may occur in stroke patients due to cerebral infarctions, symptoms such as dizziness, nausea, the sensation of the surroundings rotating around them and balance impairments are sometimes reported. This is implemented by moving the camera additionally to the participant's head movement. The anomaly is parameterized by specifying the amount and velocity of longitude and latitude movement.

4. PROCEDURE

The whole test took about 90 minutes per person. The main parts were:

Before the session; Information was given about the test, its general aim, equipment and purpose. Just before the session, the test leader and the participant made sure that there were no misconceptions about the term empathy and agreed on the definition being "empathy for another person's experience". The user learned to navigate the wheelchair. After this the sensors were connected and the HMD were adjusted comfortably and information about the three tasks was given.

During the session; the participants got time to familiarize with the VE and the supporting technologies. No time limits were set for the different tasks but a total test time limit of 20 minutes was set in order to reduce cyber sickness. Observations of real and virtual behaviour were made. Directly after the session, recorded interviews took place.

5. DATA COLLECTION AND ANALYSIS

To get more sources for interpretation the subjects were observed during their interaction with the virtual world and trying to understand the meaning of the subjects' virtual and real actions and interactions. We wanted to generate hypotheses using a version of Grounded Theory (Strauss et al. 1990). To examine the human – machine interactions we wanted to be able generate theories and hypotheses using a version of Grounded Theory. Grounded Theory has in this study been used as source of inspiration and whose characteristics are explorative, unconditional, aims to generate theories and hypotheses. The design was explorative with a semi usability character. The main approach was to view the participants as a human base of knowledge that could be an active part in exploring the emotional impact of interaction with the stroke simulator and its usability. The test design made from a non-performance perspective i.e. the participant couldn't perform well or bad, just perform.

The interviews consisted of two parts the first aimed to gather spontaneous information from an emotional response perspective and the second part was based on question based usability dialog.

6. RESULTS AND CONCLUSIONS

Results from observations and interviews indicate that the simulator in spite small problems of usability were effective in influencing caregivers empathy. All tested individuals expressed some sort of stroke, and empathy like experiences. Indicating that the anomalies developed in close collaboration with stroke affected individuals may in some way resemble a stroke experience. Considering empathy and the vast different terms and definitions historically made, we have found correlations between observed behaviours, information from interviews with different parts of the emphatic process. Observed behaviours and responses were; acting like a stroke patient in the environment such as not using their left hand for navigation and interaction even though it was working. Frustration and anger was seen during the observations resulting from experiences of not being able to cope with the everyday tasks available in the VE. Typical stroke behaviours, such as problems navigating in the hospital environment were seen among the caregivers in the virtual apartment while these professionals were performing the various tasks. Problems with orientation and dizziness, commonly found among stroke patients were also evident among the caregivers while in VE. All these factors, probably contributed to added knowledge and understanding among the participants in their role as clinical caregivers and new perspectives of their understanding for their patients. Still the exact mechanisms involved in gaining these new perspectives remains to be elucidated, and calls for further research in this field.

Psychological phenomena related to empathy were also evident, such as the overtaking of patient role, activation of cognitive schemes concerning stroke patients and the triggering of old traumatic experiences. This seems to indicate that most participants experience different variations of the whole process of empathy. More research is needed to make adequate measurements on what specific aspects that are influenced.

As medicine today are rapidly progressing there is a need for professional continuous education. This professional education is needed both for the development of functional rehabilitation teams around stroke patients and their families, and to be able to take today's knowledge a step further. Here we have presented a system now tested by professional caregivers who are normally working on a daily basis with stroke related issues. These results are encouraging since it appears that that this stroke simulator is usable for training caregiver's empathy for stroke patients, possibly creating an increased understanding for stroke patients daily problems. The results found have initiated studies concerning caregiver-VR-training, and the tentative clinical use of this system. Validation and evaluation of the anomalies, technical development and emotional impact are other aspects of ongoing research in our departments.

6. REFERENCES

R M Banos, C Botella, C Perpina, M Alcaniz, JA Lozano, J Osma and M Gallardo (2002), Virtual reality treatment of flying phobia, IEEE *Trans Inf Technol Biomed*, Sep; 6(3):206-12.

P M Emmelkamp, M Krijn, A M Hulsbosch, S de Vries, MJ Schuemie and C A van der Mast (2002), Virtual reality treatment versus exposure in vivo: a comparative evaluation in acrophobia, *Behav Res Ther*. May; 40(5):509-16.

J Feng, J Lazar and J Preece (2004), Empathy and online interpersonal trust: a fragile relationship. *Behavior & Information Technology,* (accepted in press).

A Garcia-Palacios, H Hoffman, A Carlin and T A 3rd Furness, C Botella (2002), Virtual reality in the treatment of spider phobia: a controlled study, *Behav Res Ther*. Sep; 40(9):983-93.

O Ghatnekar, U Persson, EL Glader and A Terént (2004), The cost of stroke in Sweden –An incidence estimate, *International Journal of Technology Assessment in Health Care,* (in press).

U Holm (1995), *Det räcker inte att vara snäll: förhållningssätt, empati och psykologiska strategier hos läkare och andra professionella hjälpare,* Natur och Kultur, Stockholm.

U Holm (2001), *Empati: att förstå andra människors känslor*, Upp, 2. Natur och Kultur, Stockholm.

J Håkansson (2003), *Exploring the phenomenon of empathy*, Department of psychology, Stockholm University.

E L LaMonica, D K Carew, A E Winder, A M Haase and K H Blanchard (1976) Empathy training as the major thrust of a staff development program, *Nurs Res*. Nov-Dec;25(6):447-51.

J H Lee, J Ku, W Cho, W Y Hahn, I Y Kim, S M Lee, Y Kang, DY Kim, T Yu, B K Wiederhold, M D Wiederhold and S I Kim (2003) A virtual reality system for the assessment and rehabilitation of the activities of daily living, *Cyberpsychol Behav*, Aug; 6(4):3838.

C R Rogers (1975), Empathic: an unappreciated way of being, *Counseling psychologist*, 5, 2-10.

S Roy, E Klinger, P Legeron, F Lauer, I Chemin and P Nugues (2003) Definition of a VR-based protocol to treat social phobia, *Cyberpsychol Behav*, Aug; 6(4):411-20.

C Sas and G O´hare (2003) The presence equation: An investigation into cognitive factors underlying presence, Teleoperators and Virtual Environments, Volume, 12, Issue 5, October , MIT Press.

Strauss, Anselm, Corbin and Juliet (1990) *Basics of Qualitative Research: Grounded Theory*

Procedures and Techniques, SAGE Publications, Newbury Park.

F Vincelli, L Anolli, S Bouchard, B K Wiederhold, V Zurloni and G Riva (2003), Experiential cognitive therapy in the treatment of panic disorders with agoraphobia: a controlled study. *Cyberpsychol Behav*. Jun; 6(3):321-8.

J Wald and S Taylor (2000) Efficacy of virtual reality exposure therapy to treat driving phobia: a case report, *J Behav Ther Exp Psychiatry*, 2000 Sep-Dec; 31(3-4):249-57.

B K Wiederhold, D P Jang, R G Gevirtz, S I Kim, I Y Kim and M D Wiederhold, (2002) The treatment of fear of flying: a controlled study of imaginal and virtual reality graded exposure therapy, IEEE *Trans Inf Technol Biomed*, Sep; 6(3):218-23.

Internet

The National Board of Health and Welfare (2004) Swedish National Guidelines for the Management of Stroke, http://www.sos.se

Technology

http://www.ascension-tech.com

http://www.fakespace.com

http://www.sgi.com

http://www.sgi.com/software/performer

http://www.isl.uiuc.edu/software/software.html

ICDVRAT 2004

Session VIII. Speech and Communication

Chair: Jaime Sánchez

Interactive rehabilitation software for treating patients with aphasia

C Sik Lányi[1], E Bacsa[2], R Mátrai[3], Z Kosztyán[4] and I Pataky [5]

[1,2,3,4]Department of Image Processing and Neurocomputing, University of Veszprém
Egyetem u. 10, Veszprém, H-8200, HUNGARY

[5]National Centre of Brain Vein Diseases OPNI,
Hűvösvölgyi u 116., Budapest H-1021, HUNGARY

[1]lanyi@almos.vein.hu, [2]erzsebet.bacsa@freemail.hu, [3]ritka@primposta.hu, [4]kzst@vision.vein.hu,
[5]pataky@opni.hu

[1]www.knt.vein.hu

ABSTRACT

Aphasia is an impairment of language, affecting the production or comprehension of speech and the ability to read or write. Most common cause of aphasia is – about 23–40 % of stroke survivors - acquired aphasia. The rehabilitation of aphasia is a medical, special treatment (speech therapy), which is the task of a psychologist. It needs long and intensive therapy. More detailed information about therapy can be found in (Engl at al, 1990, Subosits, 1986). In this paper we present our implementation or realization of interactive multimedia educational software to develop readiness of speech for helping the therapy. The software were developed within the frame of youth scientific and MSc thesis works. The first program was developed in Flash, the second in Macromedia Director. The goal of our software is to teach the most important everyday words. The software will be a useful device in the education of children with heavy mental deficiencies. Reading the program you can learn how it works and what current results we have achieved.

1. WHAT IS APHASIA?

Aphasia is an impairment of language, an acquired communication disorder that impairs a person's ability to process language, but does not affect intelligence. It also impairs ability to speak and understand others and most people with aphasia experience difficulty in reading and writing. While aphasia is most common among older people, it can occur in people of all ages, races, nationalities and gender. Further information about Aphasia can be found in (Leel and Ősi, 1988).

1.1 Types of Aphasia

Global Aphasia is the most severe form. Patients with this type of aphasia produce only few recognizable words, can understand little or no spoken speech at all and they can neither read nor write. These symptoms can usually be seen after a patient has suffered a stroke and they may rapidly improve if the damage has not been too extensive and there has not been greater brain damage or more severe and lasting disability.

Broca's Aphasia means that speech output is severely reduced, limited mainly to short utterances of less than four words. As a result vocabulary access is limited, the formation of the sounds is often laborious and clumsy. Patients with Broca's aphasia may understand speech and be able to read but the quality of speech is halting and effortful. Furthermore, they are very limited in writing.

Mixed non fluent aphasia means sparse and effortful speech resembling Broca's aphasia. Patients of this illness are limited in comprehension of speech and they do not read or write beyond elementary level.

In case of Anomic Aphasia persons are left with a persistent inability to supply words for the things that they want to talk about. Mainly they have problems with nouns and verbs. However they can understand speech well and in most cases they can read adequately. Unfortunately they have poor writing ability.

1.2 What causes Aphasia?

Most common cause of aphasia – about 23 – 40 % of stroke survivors acquire aphasia. It can also result from head injury, brain tumour or other neurological causes. It is estimated that about one million people in the United States have acquired aphasia, or 1 in every 250 people.

More common than Parkinson's Disease, cerebral palsy or muscular dystrophy. About 1/3rd of severely head-injured persons have aphasia. However, most people have never heard of it.

1.3 Recovery from Aphasia

After stroke – if symptoms last longer than two or three months, complete recovery is unlikely. However, it is important to note that some people continue to improve over a period of years and even decades. Improvement is a slow process that usually involves both helping the individual and the family understand the nature of aphasia and learning compensatory strategies for communicating.

How do you communicate with a person with aphasia? Communication Do's and Don'ts.

First and foremost, a person with aphasia has to be considered as a normal human being, so talk to them as an adult not as to a child. Minimizing or eliminating background noises helps a lot in understanding. Before starting communication make sure you have the person's attention. To make them feel confident encourage them all the time. Use all modes of communication: speech, writing, drawing and yes-or-no responses to avoid being boring during studying. Give them enough time to talk and permit a reasonable amount of time to respond as well. It is essential that you should accept all communication attempts. Keep your own communication simple but adult-like. It is better to use simplified sentence structures and keep in mind to reduce your rate of speech. Keep your voice at a normal level and do emphasize key words. If it is possible, use a lot of gestures and aids during communication. Do not hesitate to repeat any statement when it is necessary but never ever attempt to finish the patients' statements for them.

2. SOFTWARE DEVELOPMENT

2.1 Developing environment

The first program was developed in Flash MX, the second in Macromedia Director.

The Flash MX has more useful tools, which helped programming the subtasks. The Flash MX is suitable to develop interactive multimedia programs, too, in contrast with other programming languages. As a matter of fact, the Flash MX is an animated software which supports the vector graphical and "raster" based pictures. This is a development environment constructing interactive programs by means of its ability to be programmed. With the help of the animated software called Flash MX interactive programs including extremely fast and spectacular multimedia elements can be created. Further information about Macromedia Flash MX can be found in (Macromedia website).

Macromedia Director is a widely used multimedia developing environment, which combines the effectiveness of high level programming languages with the show of animation studios. The developing of interactive multimedia programs is complicated and takes a lot of time with conventional programming languages, such as Pascal or C++, but with the help of Director these applications can be written more easily and quickly. Further information about Macromedia Director can be found in (Macromedia website).

We will make a comparison between of the two software so that the users can decide which one has more advantages for them and is worth further development in an other article. The comparison and analysis of developing environments is not the theme of this article. The major characteristic feature of the program developed in Flash MX is its extraordinary speed that can be achieved with not using much memory. Due to the low number of Kbytes used, it can be used through the Internet very easily. What helps the Aphasic patients the most is that they do not need to study in rehabilitation centres but owing to the flexibility of the program they can also practise in their homes. Though they need the guidance of a teacher in every case. The program developed in Director is used only at home or in the rehabilitation centre. It is not downloaded from the Internet.

2.1 Progress of development

First of all, we got acquainted with the topic through several books special teachers consultants offered us. Naturally, after immersing in studying the disease itself we consulted with the therapists in the school a couple of times. Their confirmations meant an essential step during the process of developing of the program since from their feedback we knew that the main idea of the program functioned smoothly, it just needed some polishing at certain parts. The basis of the multimedia software are the books “Everyday tasks” by Jacqueline Stark and “Speak without fear” by Takácsné Csór Marianna. The second book has more exercises becoming more and more difficult through the whole book. After collecting the tasks into several groups we designed the main menu. We took the following items into account: simplicity, easy to survey, easy to use, do not distract the attention.

The programming was started after the searching for simple figures and illustrations easy to survey. At the beginning we had to consider that there were a lot of similar exercises. The program had to be terminable. The number of the visible pictures on the screen is not the same in every task therefore we had to watch out for their sizes. The order of the cast members on the stage was also important. The pictures are randomly arranged in every skill level of the program. In this way protection against monotony is assured. By using the interactivity of our software the computer informs the user about the right or wrong solution. Both cases are marked with adequate figure and sound.

3. THE INTRODUCTION OF THE SOFTWARE

The application of an educational software can have difficulties in Hungarian schools and rehabilitation institutes because of the low number and quality of the computers. Unfortunately the situation is the same with the equipment of the psychologist and special teachers who take care of the aphasia patients. It was an important point of view that the programs should also run easily on an average computer to avoid any difficulties. Considering these factors we have chosen programming on HTML basis because of its being modular and having a small size. The Microsoft Windows operating system is generally installed on school computers. The objects and their interactivities were made with Flash MX, which does not need a compiler and the building of large executable files, only the downloading of the Flash Player 6.

We can direct the interactive objects of the film in the Director, as a fictive director. The action is taking place on the stage and it is visible for the user. The cast members are pictures, video clips, music and all of the buttons and tools. Events can be added to the objects and with them we can handle the information received by the computer.

The program starts in both cases with choosing one of the tasks. The user may not be able to use the software alone in this phase of the rehabilitation, so a teacher or a therapist’s guidance is inevitable. Answering the questions does not happen in a fixed order, so it is possible to begin the practice from his (or her) own skill level.

We take care of the interactive participation using the possibilities of the manipulation technique in every task. The number of the right answers isn’t displayed with points but only with illustrating animations. The programs are user friendly, they do not require special knowledge and everybody is able to acquire how to use them in a few seconds.

Requirements: Pentium class or equivalent CPU, SVGA display with 800*600 resolution, Windows operating system with included Internet Explorer WEB browser, Mouse, Sound card, Macromedia Flash Player 6.

3.1 The structure of the menus

There are three pictures in the main menu of the first program (Figure 1). Here the patient or the teacher can choose from the language with a mouse click. The next step after choosing the language is giving the user’s first and family name. User’s name is required for the further developing of the records.

You can get into the next level only if you fill in the field of the User’s name. Then you can choose from the given rooms. (Figure 3) The second program shows only the objects of the house. The main menu of the second program starts with a user’s guide. (Figure 4)

4. THE TASKS

The main task of both pieces of software is to teach the everyday words of the house or of the environment near the house. (Figure 5) In every submenu there are 4 skill levels.

The first level makes the user acquainted with the objects of the room. (Figures 6 and 7) If the user points with the mouse to an object, the name of the object appears on the screen and it can be read. This first skill level trains the user’s memory power.

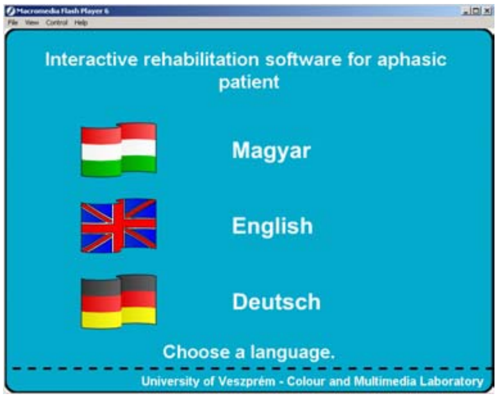


Figure 1. *Main menu (Flash version).*

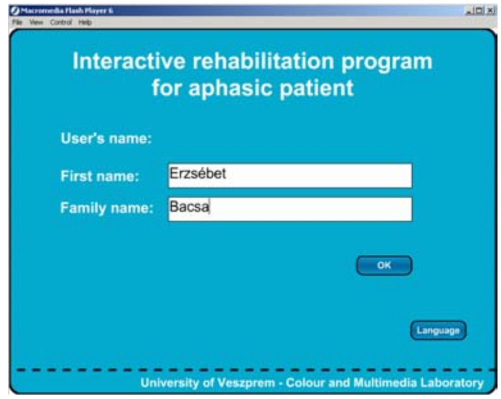


Figure 2. *Submenu - User's name*

Figure 3. *Submenu – rooms (Flash version).*

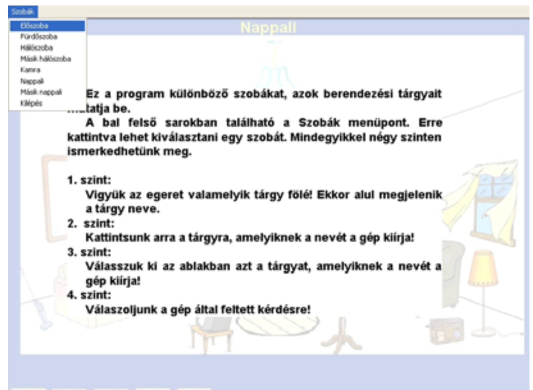


Figure 4. *Main menu (Director version)*

Figure 5. *Menu of room (Flash version)*

The second level shows an uncoloured picture. The program asks the user to show the named objects. (Figures 8 and 9) If the user manages to find the asked object, the object turns into a coloured one. In this way the user can make the whole picture colourful. If the user is unsuccessful, the program sends an error message and the user can continue.

The user's ability is also controlled on the third level. The earlier learned room is in the background, but it is pale. (Figures 10 and 11) The program shows three figures randomly and the user has to choose one with answering a question. This is essential because the user cannot recall its name but only its place properly. The real knowledge is examined at this level.

If the user could not answer, the program sends an error picture. This third level asks questions as long as the user or the teacher does not choose another skill level. The fourth level asks more difficult questions than the third level. The questions deal with the handling of the objects. (Figures 12 and 13)

Figure 6. *1st skill level (Flash).*

Figure 7. *1st skill level (Director)*

Figure 8. *2nd skill level (Flash).*

Figure 9. *2nd skill level (Director)*

Figure 10. *3rd skill level (Flash).*

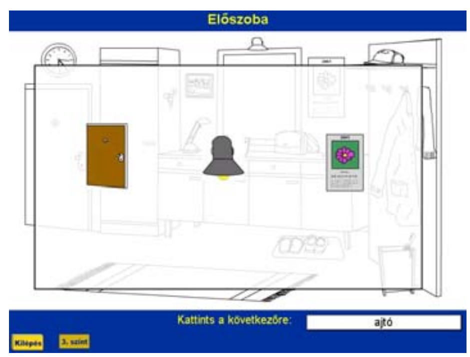


Figure 11. *3rd skill level (Director)*

Figure 12. *4th skill level (Flash).*

Figure 13. *4th skill level (Director)*

If the user can find the object, the program asks its name and the user has to write it into an input row on the screen. After the right answer the program asks a new question. The therapist chooses the starting level and the tasks. It can be used with the therapist's assistance.

5. SUMMARY

Within the frame of youth scientific and MSc thesis works we have prepared interactive multimedia educational pieces of software to develop the readiness of speech for helping the therapies. The software package contains two programs. The first program was developed in Flash, the second in Macromedia Director. The goal of our software is to teach the most important everyday words. We made the first tests in the Kozmutza Flora Special Primary School and Talent Developing / Training Collage. We have tested only the user interface design so far. In the second step during second term in the 2003/2004 academic year we are going to make the efficiency examination not only in this school but in the National Centre of Brain Vein Diseases too.

People with difficulties in naming objects have been tested up to now. They managed to get through levels 1-3 easily. At the 4th level naming of the objects, especially in case of long words, required more time than at the previous ones. The aim of testing is to be able to use the learned material in real life. One result of the first test is that the software is a useful device not only in the education of aphasic but heavy mental deficient children too. Moreover, both the children and the teachers can use it easily. Considering the advice of the teachers helping us during the program development, the children need some motivating animation reassuring the right answer and increasing their interest for the oncoming items. On the contrary, adults do not require such motivating exercises or tasks. The English and German versions and the audio confirmation of the words are being developed at present.

Acknowledgements: Hereby, the authors, express their thanks for the enormous help they received from the teachers of the Kozmutza Flora Special Primary School and Talent Developing / Training Collage in Veszprem. The authors would like to thank for the assistance of NKFP 2/052/2001 project.

8. REFERENCES

K Cseh, Á Hegyi (1995), *Exercises for the cognitive therapy of Aphasia, (Gyakorlatok az afázia kognitív nyelvi terápiájához),* Nemzeti Tankönyvkiadó, Bp.

E M Engl, A Kotten, J Ohlendorf, E Poser (1990), *Exercises for the Aphasia Therapy,(Gyakorlatok az afázia terápiájához)*, Medician, Bp.

Á Hegyi (1995), *Aphasia Therapies: Proposal for the cognitive lingual therapy of Aphasia, (Afáziaterápiák: Javaslat az afázia kognitív nyelvi terápiájára),* Nemzeti Tankönyvkiadó, Bp.

S Jacqueline (1992), *Everyday tasks Handbook, (Mindennapi tevékenységek kézikönyv)*, CIE, Wien,

L Leel – Őssy (1988), *The paralysed man, (A megbénult ember)*, Statisztikai Kiadó, Bp.

I Subosits (1986), *Collection book of selected passages of the clinical disturbance in speaking, (Szemelvénygyűjtemény a klinikai beszédzavarok köréből)l*, Tankönyvkiadó Bp.

M Takácsné Csór (1986), *Speak without fear! Exercise Book for the therapy of Aphasia, (Beszélj bátran Feladatgyűjtemény az afázia terápiához)*, Tankönyvkiadó Bp.

The National Aphasia Association's website *www.aphasia.org*

The national Macromedia website *www.macromedia.com*

Real-time clarification filter of a dysphonic speech and its evaluation by listening experiments

H Sawada[1], N Takeuchi[2] and A Hisada[3]

[1,2,3] Department of Intelligent Mechanical Systems Engineering, Faculty of Engineering, Kagawa University
2217-20, Hayashi-cho, Takamatsu-city, Kagawa, 761-0369, JAPAN

sawada@eng.kagawa-u.ac.jp

[1]www.eng.kagawa-u.ac.jp/~sawada

ABSTRACT

This paper presents a digital filtering algorithm which clarifies dysphonic speech with the speaker's individuality preserved. The study deals with the clarification of oesophageal speech and the speech of patients with cerebral palsy, and the filtering ability is being evaluated by listening experiments. Over 20,000 patients are currently suffered from laryngeal cancer in Japan, and the only treatment for the terminal symptoms requires the removal of the larynx including vocal cords. The authors are developing a clarification filtering algorithm of oesophageal speech, and the primal algorithm of software clarification and its effectiveness was reported in the previous ICDVRAT. Several algorithms for the clarification have been newly developed and implemented, and are being evaluated by questionnaires. The algorithms were extended and applied for the clarification of the speech by the patients of cerebral palsy.

1. INTRODUCTION

A voice is the most important and effective medium employed not only in the daily communication but also in logical discussions. Only humans are able to use words as means of verbal communication, although almost all animals have voices. Vocal sounds are generated by the relevant operations of the vocal organs such as a lung, trachea, vocal cords, vocal tract, tongue and muscles. The airflow from the lung causes a vocal cord vibration to generate a source sound, then the glottal wave is led to the vocal tract, which works as a sound filter as to form the spectrum envelope of a particular voice. If any part of the vocal organs is injured or disabled, we may be involved in the impediment in the vocalization and, in the worst case, we may lose our voices.

Over 20,000 patients are currently suffered from laryngeal cancer in Japan, and the only treatment for the terminal symptoms is to remove the larynx. The removal of vocal cords means the loss of the voice, and causes various difficulties in the communication with other people, since the employment of a voice is essentially important for humans to make verbal communications.

There are mainly two ways to recover voice. One is to use an artificial larynx, which is a hand-held device with a pulse generator that produces a vocal cord-like vibration. An electrolarynx has a vibrating plastic diaphragm, which is placed against the neck during the speech. The vibration of the diaphragm generates a source sound in the throat, and the speaker then articulates with the tongue, palate, throat and lips as he does for the usual vocalization. The device has an advantage to be used by just being held to the neck and to be easily mastered, but the sound quality is rather electronic and artificial. Furthermore one hand is occupied to hold the device during the speech, which disturbs the gestural communication.

The other way is to train oesophageal speech, which is a method of speech production using an oesophagus (Sato, 1993; Max *et al*, 1996). In the speech, air is inhaled and caught in the upper oesophagus instead of being swallowed, and then the released air generates the oesophagus vibration to produce a "belch-like" sound that can be shaped into speech. A patient has difficulties to master the oesophageal speech (several years are ordinary required for the practice), however the voice is able to keep the speaker's individuality since the speech is generated by his own vocal organs, although several distinctive characteristics exist. Moreover the speaker is able to employ his body parts such as hands and facial expressions freely and actively for the communication to assist the speech.

Cerebral palsy (CP) is a condition caused by an injury to the parts of the brain which controls the ability to use muscles and bodies. The injury may happen before birth, sometimes during delivery, or soon after

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

being born. Severe CP may affect plural parts of patient's physical abilities, and requires to use a wheelchair and other special equipments to move. However, most of the patients with CP doesn't have mental retardation or disorder. CP doesn't get worse over time, and most children with CP have a normal life span. Patients of CP often have difficulties in moving and controlling vocal apparatus due to the insufficiency of the muscle controls, and the clarity of vocalized sounds is low.

The authors are developing a clarification filtering algorithm of oesophageal speech, and the primal algorithm of software clarification and its effectiveness was reported in Hisada and Sawada (2002). Several algorithms for the clarification have been newly developed and implemented, which were evaluated by questionnaires. The algorithms were extended and applied for the clarification of the speech by the CP patients. The paper presents the recent developments of the clarification filtering for the dysphonic speech, together with the implementation to a PC to be tested by a patient for the practical use.

2. BACKGROUND OF STUDIES CONCERNING DYSPHONIC SPEECH

Several researches which analyze the characteristics of the oesophageal speech have been reported so far (Noguchi and Matsui, 1996; Doi *et al*, 1996), and a device to improve the oesophageal voice is now commercially available (*e.g. VivaVoice* by Ashida Sound Co., Ltd.). The device is a package of a small circuit board equipped with a compact microphone and speaker, and transforms an unclear voice into clear by using the formant synthesis techniques, which has the disadvantage in keeping the speaker's individuality in the voices. Acoustic characteristics of the oesophageal voice have been also studied (Lu *et al*, 1997; Bellandese *et al*, 2001; Robbins *et al*,, 1984), and have accounted for the lack of the phonetic clarity. There still remain unknown characteristics and features in the oesophageal voices, and the clarification theory and algorithm are not yet established.

Patients of CP, on the other hand, often have difficulties in moving and controlling vocal apparatus, and the clarity of vocalized sounds is low. To assist the verbal communication, many support tools such as *Talking-aid* and *Kinex* have been commercially available (Koroko, 2003). However, no studies to clarify patient's voice in real-time are not found so far.

3. PURPOSE OF THE STUDY

The purpose of the study is to develop clarification filtering algorithms of oesophageal speech and the speech of CP, for presenting a hand-held device to clarify the speech in real-time with speakers' individuality preserved. Spectra of an oesophageal voice and a CP voice are shown in Figure 1, together with the comparison of a laryngeal vocalization.

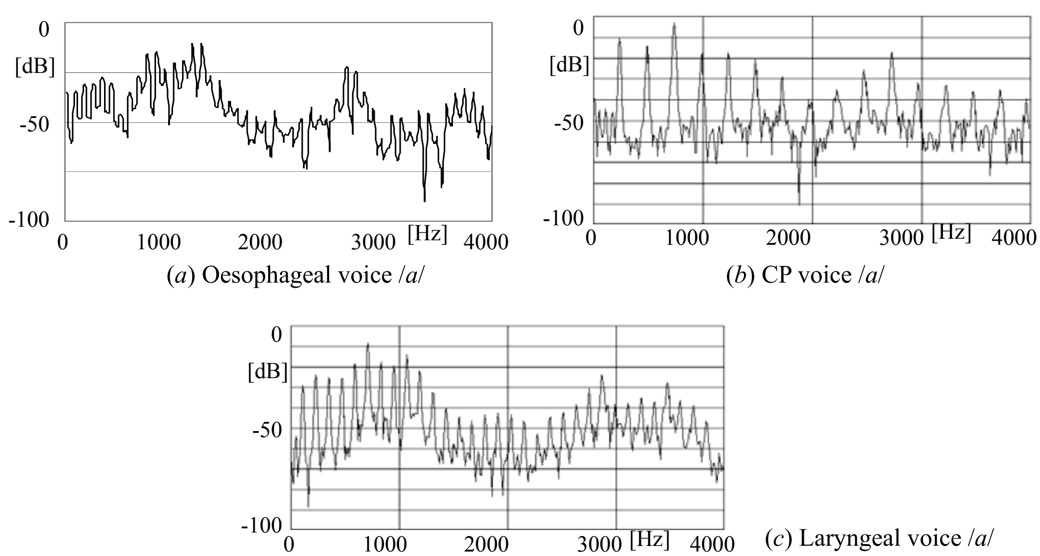


Figure 1. *Comparison of spectra of three speeches*

Typical oesophageal speech is characterized as follows.

ES1) The voice contains noises caused by the irregular vibrations of oesophagus.

ES2) Fundamental frequency is rather low.
ES3) Fundamental frequency and its overtone structure are not clear.
ES4) Fluctuation exists, which causes the instability of the voice.
ES5) Spectrum envelope is flatter than the laryngeal voice.
ES6) Volume is low because of the shortage of the expiration. (Average is 150 ml; laryngeal voice is 2000ml)
ES7) Mastering the oesophageal speech requires repetitious training, and the progress is slow and is hard to be recognized by the patient in training.

The features of a CP voice, on the other hand, are listed below:

CP1) Spectrum envelope is flatter than laryngeal voice.
CP2) Formant frequency is unstable due to the fluctuation.
CP3) Starting of utterance is indistinct and the consonants are vague, due to the shortage of the expiration.
CP4) Fundamental frequency is rather high, and its overtone structure is indistinct in the higher frequency range.
CP5) Distinctive resonance characteristic is found around 2 kHz.
CP6) Spectrum power around 3 kHz decreases.

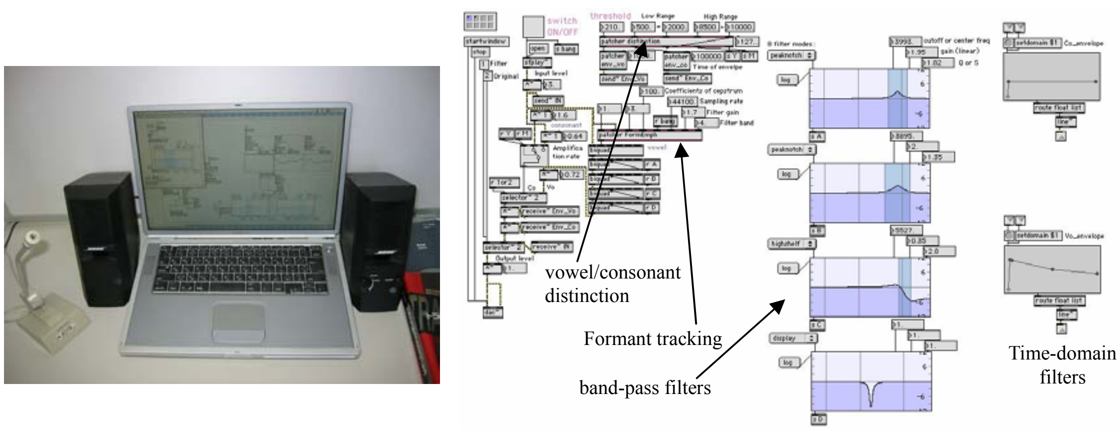
In this study, a real-time software filter for the clarification of the oesophageal speech and cerebral palsy speech has been constructed by taking account of the characteristics listed above. The implemented algorithms were tested by inputting actual dysphonic speeches of oesophagus vocalization and CP patients, and the filtering ability was evaluated by an listening experiment.

4. FILTERING ALGORITHMS

4.1 System Configuration

The algorithm was realized with the Max/MSP programming environment (Roads, 1996) in Macintosh Power Book (Power PC G4/667 MHz), equipped with a microphone and a pair of stereo speakers as shown in Figure 2(*a*). Sampling frequency was set to 44.1 kHz with 16 bits input, and the real-time processing is conducted in the programs. The software filter constructed in the Max/MSP is also shown in Figure 2(*b*).

Figure 3 shows the flow diagram of the filtering. First the input signal is judged to distinguish vowels and consonants, and is divided into two procedures. For a vowel sound, fundamental frequency is extracted to form a comb filter as to enhance pitch components, and at the same time two peaks are extracted from the spectrum envelope to be enhanced. For a CP voice, due to the features CP5) and 6), only the extraction of the first formant is executed to enhance the clarity of the vowel sounds without changing the speaker's individuality. The consonant sound, on the other hand, is enhanced in the time domain by multiplying its amplitude by a suitable coefficient. The calculations above are processed in real-time for every input to generate clarified speech outputs.



(*a*) System configuration (*b*) Clarification filter in Max/MSP environment

Figure 2. *Flow diagram of filtering.*

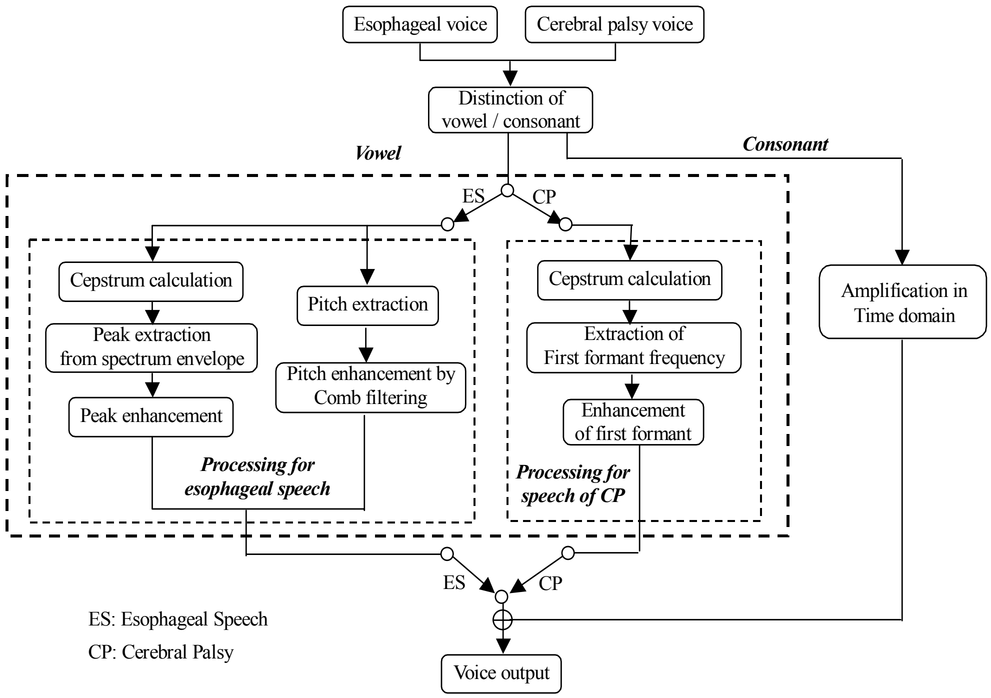


Figure 3. *Flow diagram of filtering.*

4.2 Distinction between vowels and consonants

Comb filter is able to enhance pitch components, and works to clarify vowel sounds in the oesophageal speech. Since the consonant sound basically doesn't have the overtone structure, the comb filtering performs inadequately. Because of the different characteristics of phonemes to be clarified, vowels and consonants are required to be classified in advance to employ different filtering techniques. We paid attention to the differences of the spectrum envelope observed in the frequency domain, and have proposed a filter to distinguish with each other.

The spectrum envelope of the vowel sound shows several local peaks in the lower frequency range, which are characterized as the formants, and the power gradually decreases as the frequency rises. The consonant sounds, on the other hand, are characterized as noises, and their spectra are uniformly distributed along the frequency range without a particular peak. The above observations justify the assumption that vowels and consonants can be separated by comparing the power balance between the low frequency range and the high frequency range. We defined an index B_{index} as

$$B_{index} = \frac{\textit{Average power of low frequency range}}{\textit{Average power of high frequency range}}$$

where the low frequency range was set to 200Hz ~ 4kHz and the high frequency range was 6kHz ~ 10kHz. For vowel sounds, the index value becomes larger because of the concentration of the power in the lower frequency range, on the other hand, the value becomes around 1 for consonant sounds. The threshold for the distinction was set to 1.5 by the experiment, which was not influenced by the amplitude of the input signal.

4.3 Comb filtering

A digital comb filter shown in Figure 4 enhances a fundamental frequency and its harmony components of input sounds. By suitably operating comb filters with the control of the parameters M, $g1$ and $g2$, oesophageal voices can be clarified with their overtone structure enhanced and their noise components controlled. The importance is to extract the fundamental frequency precisely. In this system we adopted a short time zero-cross method for the pitch calculation. The extracted pitch works to form a comb filter with its pitch span and enhancement ratio suitably defined. The individuality of a speaker can be preserved, since the comb filter enhances the overtones without changing the formants and the spectrum envelope.

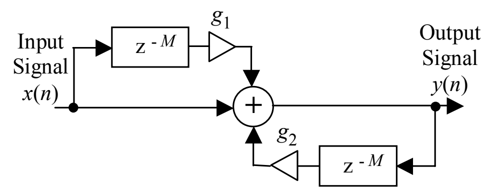
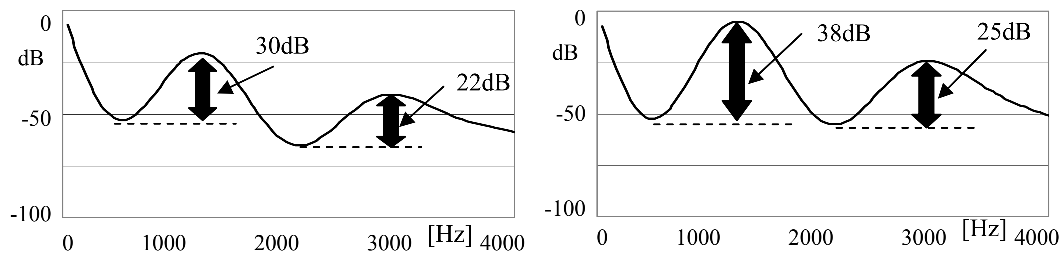
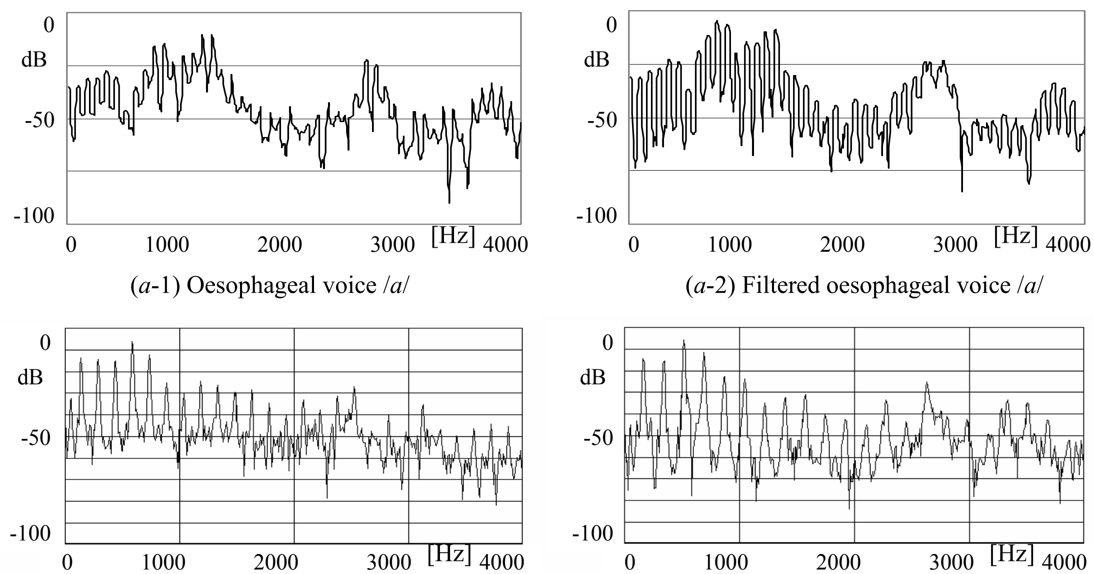


Figure 4. *Digital comb filter*



(*a*) Before enhancement (*b*) After enhancement

Figure 5. *Spectrum envelope of /a/ sound before and after the enhancement*



(*a*-1) Oesophageal voice /*a*/ (*a*-2) Filtered oesophageal voice /*a*/

(*b*-1) CP voice /*a*/ (*b*-2) Filtered CP voice /*a*/

Figure 6. *Results of clarification filtering for CP voice /a/.*

4.4 Enhancement of peaks in spectrum envelope

A spectrum envelope characterizes the resonance characteristics, which are perceived as different vowel sounds. Vowel sounds are formed and articulated according to the inner shape of the vocal tract, which is governed by the complex movements of the jaw, the tongue and the muscles. Due to the insufficient articulation, the spectrum envelope of a dysphonic voice is flat comparing with a laryngeal voice. In this study, we tried to clarify vowel sounds by enhancing the spectrum peaks and modifying the resonance characteristics.

Spectrum envelope is obtained by the inverse Fourier transform of the 1st to 32nd coefficients extracted by the cepstrum calculation. Formants are the local maxima in the spectrum envelope, and the first and the second largest maxima in the lower frequency range are extracted every 30 ms. Then two band-pass filters to enhance the two peaks are generated in real-time, and are applied to emphasize the envelope. Since the higher-order formant frequencies are unstable due to the instability of vocalization, the peak emphasis filtering is applied only to the first and second formants. Figure 5 shows the spectrum envelope of /*a*/ sound before and after the enhancement.

4.5 Consonant enhancement

The volume of the oesophageal speech is generally low because of the lack of the expiration pressure, and the consonant pronunciations are insufficient and difficult to be distinguished. The starting utterance and consonants of CP speech are also indistinct as listed in CP3). The consonant sound is enhanced in the time domain by multiplying its amplitude by a suitable coefficient. The amplitude can be varied from 1.0 to 3.0 in the program not to distort the vocal characteristics, and a suitable value is determined according to the characteristics of input voice based on listening experiments conducted before the use.

4.6 Filtering experiments

Filtering experiments for oesophageal speech and CP voices were conducted. Figure 6 (*a*) and (*b*) show the power spectra of oesophageal speech /*a*/ and CP voice /*a*/ before and after the filtering, respectively.

The proposed filter could form the clear overtone structures of the fundamental frequencies for oesophageal speeches. The basic form of a spectrum envelope and formant positions were well preserved throughout the frequency range, and the peaks and valleys of the overtones were enhanced, which contributed to the clarification of vowel sounds and the noise reduction. The starting of utterance and consonants were emphasized and clarified, owing to the enhancement of the signal in the time domain. The individuality of the speaker was satisfactorily preserved, and the improvement of the clarification was also recognized.

5. LISTENING EXPERIMENTS

For the assessment of the filtering ability, listening experiments were conducted and filtered voices were evaluated by questionnaires.

10 able-bodied subjects listened to 12 pairs of unfiltered and filtered oesophageal speeches given by 6 speakers, and evaluated them in 7 levels from 12 points of view. The contents of the voices are

Voices A_1 - F_1: Speech of Japanese 5 vowels /aiueo/ by 5 speakers A to F.
Voices A_2 - F_2: Reading of Japanese sentence /Chiisana Otokonoko ga Teeburuno ueni/ (A little boy is sitting on a table) by 5 speakers A to F.

To examine the filtering ability to the difference of training experience, the speakers A to F were selected from the experience of a half, 3, 4, 7, 17 and 18 years, respectively. Only the speaker E was a female.

For CP voices, the subjects listened to 6 pairs of unfiltered and filtered speeches given by 3 patients, and scored them in 7 levels. The contents of the voices are

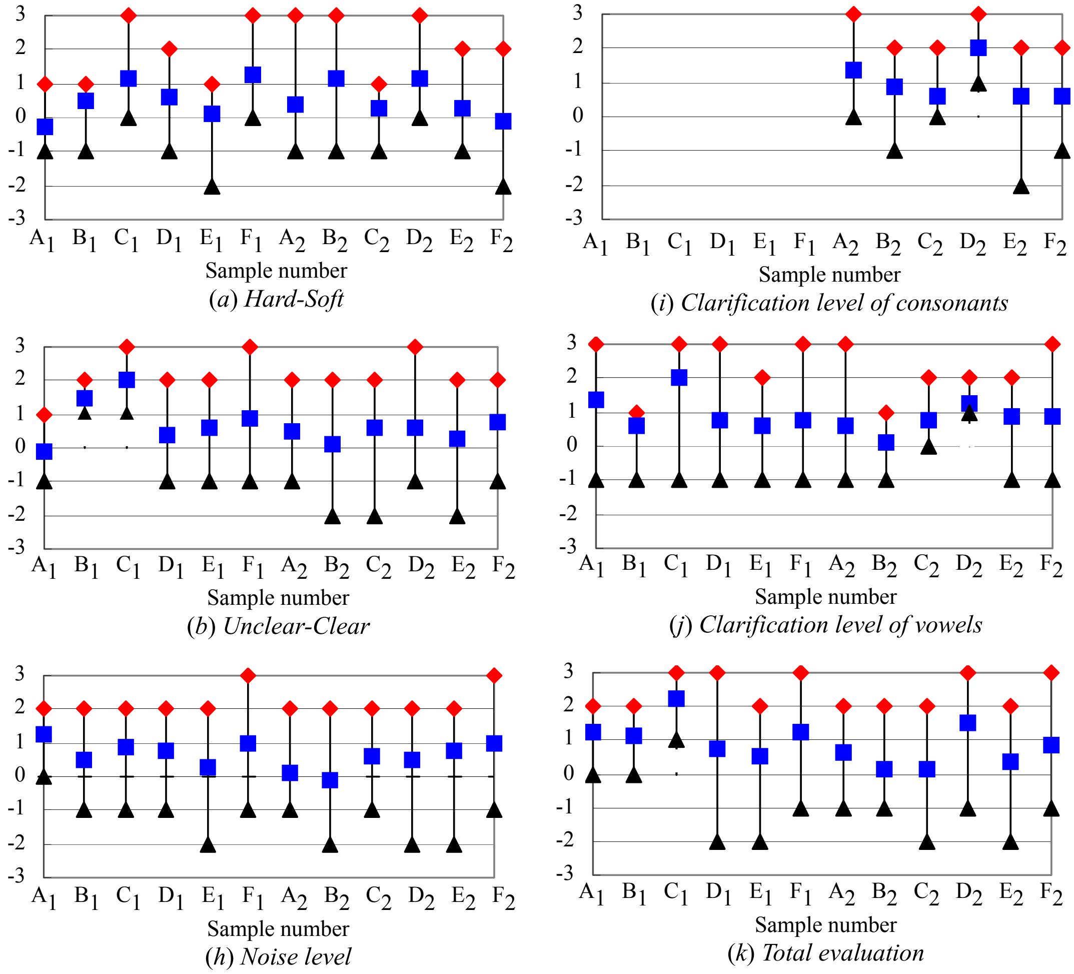
Voices G_1 - I_1: Speech of Japanese 5 vowels /aiueo/ by 3 speakers G to I
Voices G_2 - I_2: Reading of Japanese sentence /Chiisana Otokonoko ga Teeburuno ueni/ (A little boy is sitting on a table) by 3 speakers G to I

Twelve evaluation points are

(a) Hard-Soft, (b) Unclear-Clear, (c) Rough-Smooth, (d) Blur-Sharp,
(e) Unpleasant-Pleasant, (f) Echo level, (g) Electronic-Human, (h) Noise level,
(i) Clarification level of consonants, (j) Clarification level of vowels,
(k) Total evaluation, (l) Preservation of individuality.

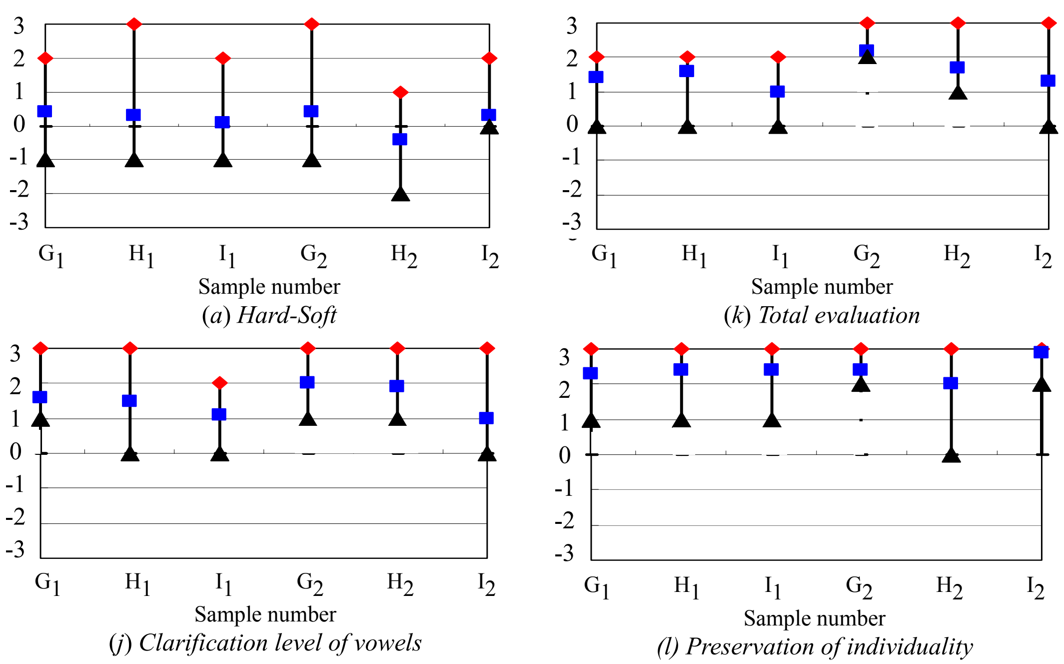
A method of pair comparisons was employed for the evaluation. First, a subject listened to a pair of unfiltered and filtered speech, and was let know which is the filtered voice. After listening, he evaluated the filtered voice by comparing with the unfiltered speech, based on the twelve evaluation points listed above. In the adjective pairs, the former was assigned as negative impression, and the latter was rated as positive impression. A subject evaluated the speech with scores between -3 and +3. The greater the points are, the higher the filtering results were better evaluated.

Figure 7 and 8 show a part of results of the listening experiments for the clarification of oesophageal speech and CP speech, respectively. The maximum scores are plotted with the mark ◆, the average scores with ■, and minimum scores with ▲, summed up by the 10 subjects' evaluations. If the average plot is allocated upper than 0 level, the filtering is considered to be effective.



(a) *Hard-Soft*
(i) *Clarification level of consonants*
(b) *Unclear-Clear*
(j) *Clarification level of vowels*
(h) *Noise level*
(k) *Total evaluation*

Figure 7. Results of listening experiment for oesophageal speech



(a) *Hard-Soft*
(k) *Total evaluation*
(j) *Clarification level of vowels*
(l) *Preservation of individuality*

Figure 8. Results of listening experiment for the speech of CP

In almost all view points, filtered voices obtained higher evaluations, especially in the points of smoothness and clearness. For oesophageal speech, the filtering was especially effective for the speakers with the training experience of 3 to 4 years. In the evaluations of ***clarification level of consonants***, ***clarification level of vowels*** and ***total evaluation***, better assessments were obtained owing to the proposed filtering algorithms.

6. CONCLUSIONS

This paper introduced a real-time software filtering algorithm which clarifies oesophageal speech and the speech of cerebral palsy patients with the speaker's individuality preserved, and presented the results of the listening experiments for the assessment of the filtering ability. The filtering results were preferably evaluated by the listeners in the points of the smoothness, clearness and the clarification of vowels and consonants, as we had designed to complement the unclear factors and to enhance the clarity. There still remain the problems such as the echo effects and electronic impressions caused by the mis-detection of voice characteristics and mis-control of the filtering parameters in the real-time processing.

We continue to construct the precise filtering algorithm to clarify the dysphonic speech with the speakers individuality preserved. We are currently working to develop a portable device using a digital signal processor (DSP) to generate naturally clarified vocalization at any time and place. The proposed software algorithm will be installed in a portable device to contribute to the use in mobile, and the study will support the verbal communication without being aware of the handicaps in speech.

Acknowledgements: This work was supported by the Grants-in-Aid for Scientific Research, the Japan Society for the Promotion of Science (No. 13780288). The authors would like to thank Dr. Yoichi Nakatsuka, the director of the Kagawa Prefectural Rehabilitation centre for the Physically Handicapped, Dr. Rieko Goto in the department of medicine, Kagawa University, and the members of Kagawa Kouyukai for their helpful supports for the experiment and the useful advice.

7. REFERENCES

T Sato, Oesophageal Speech and Rehabilitaion of the Laryngectomized Kanehara & Co., Ltd., Tokyo, 1993

L Max, W Steurs, and W De Bruyn, Vocal capacities in oesophageal and tracheoesophageal speakers Laryngoscope, 106, 93-96, 1996

A Hisada and H Sawada, Real-time Clarification of Oesophageal Speech Using a Comb Filter *International Conference on Disability, Virtual Reality and Associated Technologies*, pp.39-46, 2002

E Noguchi and K Matsui, An evaluation of oesophageal speech enhancement *The Acoustical Society of Japan, Autumn Meeting* 2-6-13, pp. 421-422, 1996

T Doi, S Nakamura, J L Lu and K Shikano, Improvement in oesophageal speech by replacing excitation components in cepstrum domain *The Acoustical Society of Japan, Autumn Meeting* 2-4-17, pp. 253-254, 1996

J L Lu, S Nakamura and K Shikano, Study on Pitch Characteristics of Oesophageal Speech *The Acoustical Society of Japan, Spring Meeting* 2-7-19, pp. 253-254, 1997

M. Bellandese, J. Lerman, and J. Gilbert, An Acoustic Analysis of Excellent Female Oesophageal, Tracheoesophageal and Laryngeal Speakers Journal of Speech, Language and Hearing Research, 44, pp. 1315-1320, 2001

J Robbins, H Fisher, E Blom and M Singer, A comparative acoustic study of normal, oesophageal and tracheoesophageal speech production . Journal of Speech and Hearing Disorders, 49, pp. 202-210, 1984

Kokoro Resource Book 2003-2004 Kokoro Resource Book Publishing, http://www.kokoroweb.org/index.html, 2003.

C Roads, The Computer Music Tutorial The MIT Press, 1996

Time-scale modification as a speech therapy tool for children with verbal apraxia

E Coyle[1], O Donnellan[2], E Jung[3], M Meinardi[4], D Campbell[5], C MacDonaill[6] and P K Leung[7]

[1,2,3]School of Control Systems and Electrical Engineering,
[4,5]School of Languages,
Dublin Institute of Technology, Kevin Street, Dublin 8, IRELAND

[6,7]Digital Media Centre, Dublin Institute of Technology,
Aungier Street, Dublin 8, IRELAND

eugene.coyle@dit.ie, *olivia.donnellan@dit.ie*

[1,2]*www.dmc.dit.ie/ditcall*

ABSTRACT

A common suggested treatment for verbal apraxia is repetition, and the use of slow speech. The required slow speech may be attained by time-scaling ordinary-speed speech. However, when used for this purpose, the quality of the expanded speech must be of a very high quality to be of pedagogical benefit. This paper describes a new method of time-scaling based on the knowledge of speech characteristics, the relative durations of speech segments, and the variation of these durations with speaking rate. The new method achieves a high quality output making it suitable for use as a computer-assisted speech therapy tool.

1. INTRODUCTION

Verbal apraxia is a motor speech disorder. It manifests itself as an inability to consistently position the articulators for the production of speech sounds, and for sequencing those sounds into syllables or words. It is not a muscle or a cognitive disorder. Generally, the sufferer has a much better understanding of language than production of language. He/she may have the concept of the words, but has difficulty in translating from the planning area of the brain to actual words or speech.

Research into treatments for apraxia is still at an early stage, and approaches differ to some extent, however, there are some common features. The common themes among suggested therapies include a high degree of practice and repetition, and the use of slow speech. Experienced therapists report that children with apraxia need frequent repetition of sounds, sound sequences, and movement patterns in order to incorporate them and make them automatic. Indeed, a CD, Moir (2000), has already been produced for this purpose. The CD consists of a number of popular children's songs, but recorded at a slower rate, so that children with speech difficulties such as apraxia can have fun singing along, while at the same time developing their speech motor skills.

Our proposal is a variation of this concept, where poems and nursery rhymes are slowed down to any desired speed using a high-quality time-scale modification algorithm. The advantage of this is that the learner can choose a speed to suit his/her level, and can speak along with or repeat any chosen segments of the slowed-down clips at the chosen suitable speed. As the child becomes more adept, the speed can be increased, until eventually the learner will be able to repeat the segments at full speed. This provides the children with access to frequent repetition of sounds and sound sequences necessary to develop their motor skills, at a speed appropriate to the particular child's level of difficulty.

Time scale modification (TSM) refers to the process of altering the duration of an audio segment. A signal may be expanded, producing a signal of longer duration (a slower signal), or compressed, resulting in a signal of shorter duration (a faster signal). To be effectively time-scaled, the modified signal must retain all the characteristics of the original signal. In particular, the perceived pitch, speaker identity and naturalness must be maintained. Simply adjusting the playback rate of the signal will alter the duration of the signal, but will also undesirably affect the frequency contents. Of the current methods for performing TSM, many are capable of producing a good quality output. However, for the proposed Computer-Assisted Speech Therapy (CAST) tool the quality needs to be extremely good, void of any distortion or unnaturalness. Current techniques simply do not produce a sufficient degree of quality, and furthermore, the quality decreases as the scaling rate increases.

Section 2 of this paper describes the general principles of a computationally efficient time-scaling technique and discusses some of the problems time domain over-lap add (TDOLA) methods encounter. Section 3 shows how to overcome these problems by suitably pre-processing the speech signal by taking into consideration specific characteristics of natural speech. This leads to a new speech-adaptive time-scaling algorithm, which provides a high quality output, even for high modification factors.

2. TIME-SCALE MODIFICATION

There are many different time-scaling techniques already developed, such as time-domain overlap-add techniques (TDOLA), frequency-domain techniques, and parametric techniques, with advantages and disadvantages to each. A full summary of these can be found in Lawlor (1999). The TDOLA approach is best suited to periodic signals such as voiced speech, and is also the best compromise of quality and efficiency, providing a high quality output for a relatively low computational load.

A TDOLA technique, in basic terms, performs time scale modification essentially by duplicating small sections of the original signal, and adding these duplicated segments using a weighting function, i.e., to make a signal have a longer duration, individual small segments are made longer my duplicating or repeating them. The technique requires firstly segmenting the waveform into a series of overlapping frames by windowing the signal at intervals along the waveform. These frames can then be added together, but with a different amount of overlap. Time-scale expansion is achieved by creating a waveform through the recombination of frames with a reduced amount of overlap. The amount of overlap required depends upon the desired expansion. Similarly, an increased amount of overlap will result in a time-scale compressed signal. The different TDOLA techniques generally vary in the way the waveform is segmented (choice of window, where segmentation occurs etc), or how successive frames are overlapped.

The most commercially popular TDOLA algorithm is the Synchronised Overlap-Add (SOLA) algorithm, Roucus (1985), because of its low computational burden with relatively high quality output. A more recent development, by Lawlor (1999), is the Adaptive Overlap-Add (AOLA), which offers an order of magnitude saving in computational burden without compromising the output quality, making it a suitable candidate for real-time implementation, such as in a computer-assisted speech therapy (CAST) package.

2.1 The Adaptive Overlap-Add Algorithm

The AOLA algorithm works in the following manner:

- A window length of ω is chosen such that the lowest frequency component of the signal will have at least two cycles within each window.
- The frame is duplicated.
- The duplicate of the original is shifted to the right to align the peaks, figure 1(b).
- Overlap-adding the original frame and its duplicate produces a naturally expanded waveform; figure 1(c). The length of this expanded segment is $\omega.ne$, where *ne* is the natural expansion factor.
- A portion of length *st* of the input signal is taken and is concatenated with the last expanded segment; figure 1 (d)-(e). *st* varies for each iteration and is a function of ω, *ne* and *de* (desired expansion factor).

$$st = \omega \frac{(1 - ne)}{(1 - de)}$$

- The next segment to be analysed is the ω-length frame ending at the right edge of the appended *st* segment, figure 1(e). This process continues until the end of the input signal is reached.

This method has a low computational load relative to other commercial algorithms of similar quality. Another advantage is that there are no discontinuities at the frame boundaries, as can be the case in other algorithms. This is because, referring again to figure 1, the area in (c) ending in the vertical dashed line and the area ending in the vertical dashed line in (d) are exactly the same shape, so the segment *st* appended to the expanded waveform will be aligned perfectly (e).

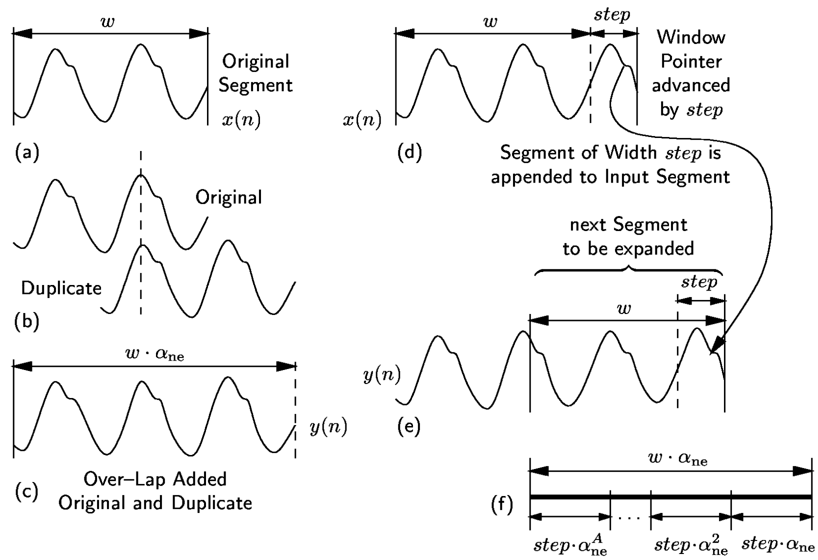


Figure 1. *AOLA algorithm, Lawlor (1999)*

2.2 Problems with Overlap-Add Techniques

Although TDOLA techniques provide the best compromise of computational load and quality, they have a few problems. As previously mentioned, TDOLA techniques perform time scale expansion by duplicating or repeating small segments of the original signal. If one of these segments to be repeated consists of a transient, such as a plosive, this may result in unwanted clicks, or what could be perceived as a 'stuttering' effect. Current TDOLA techniques assume that all segments of the speech, whether voiced, unvoiced, vowel or consonant, should be time-scaled at a uniform rate. A problem with this is that the transients (e.g. plosives in speech or drumbeats in music) are time expanded to the same degree as non-transient segments of the original signal. If a plosive were expanded using the TDOLA technique, the sound of the plosive would be distorted, and intelligibility of the resulting speech will be diminished. Also, vowels tend to be more influenced by speaking rate the consonants. To maintain intelligibility and naturalness, different time scaling factors need to be applied to the different segments of speech.

3. SPEECH-ADAPTIVE TIME-SCALE MODIFICATION

The intention of the proposed time-scaling algorithm is to be able to slow down speech while maintaining the naturalness and retaining all the characteristics of the original signal. To be capable of doing this certain aspects about the nature of the speech segments, the relative durations of these segments and the variation of these durations with speaking rate are considered for the time-scaling.

3.1 Relative variances in durations of vowels, voiced consonants and unvoiced consonants

It has been noted by Ebihara (2000) that the duration of unvoiced segments of human speech varies less than the duration of voiced segments. Ebihara recommends that a non-uniform rate be applied when time-scaling speech, so as to maintain the temporal structure of the utterance. He proposes a method of modifying only the voiced segments or alternatively only the vowel segments. Kuwabara (1997) backs up this theory by pointing out that changes in duration due to speaking rate are most obvious in voiced segments and particularly in vowels. He claims that the duration of voiced consonants varies more strongly with variations in speaking rate than that of unvoiced consonants.

By this reasoning, it can be concluded that, to imitate real speech characteristics, vowel sounds need to be more affected by time-scaling than consonants, and voiced consonants more than that of unvoiced consonants.

3.2 Consistency in duration and character of plosives

As previously mentioned, the effect of time-scaling on plosives is undesirable. As plosives convey a large amount of information, it is necessary to preserve their character under TSM. At large TSM factors, plosives may be artificially transformed into fricatives, e.g. /p/ slowed down at a high TSM factor may sound more like the fricative /f/. In normal speech, the closure stage of a plosive tends to be consistent in duration, regardless of the speed of the speech. The duration of the burst also tends to be constant and the 'suddenness' of the onset of energy needs to be maintained. Time-scaling the burst leads to transient repetition, therefore to maintain the character of plosives the closure and burst need to be directly translated to the output without applying TSM.

3.3 Algorithm flowchart and summary

The following diagram (Figure 2) summarises the procedure proposed to achieve speech adaptive time-scale modification. Firstly, each segment of the input signal is examined to verify that speech exists. If no speech exists, the segment is assumed to be silence, for example a pause between phrases or sentences, and this segment of silence will then be time-scaled with a scaling factor of α_1. If speech exists, the type of speech contained in the segment must be determined. If analysis of the segment reveals that it is a plosive, or part of a plosive (closure or burst), the segment is copied to the output without any time-scaling, so as to preserve the nature of the plosive. If the segment is not a plosive, then it contains speech that is either voiced or unvoiced. Voiced speech is further analysed to determine whether it is a vowel or voiced consonant. As vowels are most influenced by speaking rate, they are time-scaled the most, with a scaling factor of α_1. The duration of voiced consonants varies less than vowels, but more than unvoiced consonants, so voiced consonants are time-scaled with a factor of α_2 and unvoiced speech is time-scaled with a scaling factor of α_3, where

$$\alpha_1 > \alpha_2 > \alpha_3 > 1$$

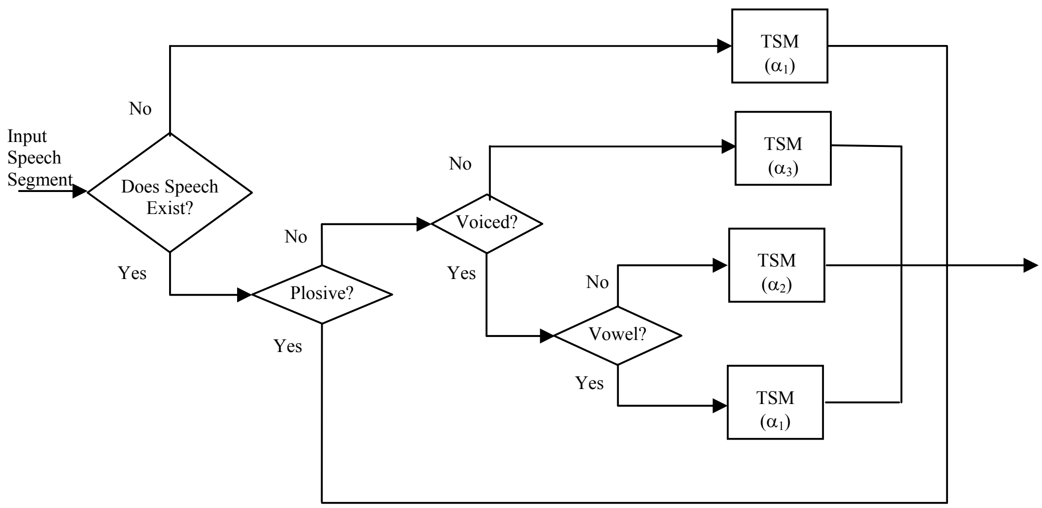


Figure 2. *Flowchart for the proposed speech adaptive time-scaling method*

4. EXPERIMENTS AND RESULTS

A number of speech samples were recorded at a sampling rate of 16 kHz, and more samples were taken from the TIMIT database. An equal number of male and female speakers were used. Each signal was then analysed on a frame-by-frame basis, and manual segmental detection was performed, i.e. through a combination of examining the waveform and listening to the signal, decisions were made as to whether the frame consisted of a plosive, vowel, voiced consonant, unvoiced consonant or silence. Matlab scripts were then written to adapt the AOLA algorithm and perform a variety of variable and uniform time-scaling methods, as described in Table 1.

For the evaluation of the performance of the different time-scaling methods a series of informal listening test was conducted. Two speech samples were recorded at a sampling rate of 16 kHz and two more samples were taken from the TIMIT database. Seven male and seven female listeners participated in the test. Each test signal was segmented depending on the utterance type (plosive, vowel, voiced consonant, unvoiced consonant or silence) and then slowed down using the methods described in Table 1. All slowing down methods were based on the AOLA algorithm and the implementation was done in Matlab.

For method D, the new proposed method, two different sets of scaling parameters were considered in order to investigate the existence of a difference in quality for different sets of parameters. For each set, the requirement of $1 < \alpha_3 < \alpha_2 < \alpha_1$ was adhered to, but the distance between the values was varied. In the first set (D1), the values varied linearly from 1 to α_1, while in the second set (D2), these values varied exponentially.

Table 1. *Compared time-scaling methods*

Method	Description
A	Uniform TSM
B	Variable TSM with voiced segments only being modified
C	Variable TSM with vowels only being modified
D	The new speech-adaptive TSM method.

4.1 Experimental Setting

All speech samples were time-scaled by each of the above methods and at three different overall time-scale modification factors, namely 2, 2.5 and 3. Two different informal listening tests were used to assess the quality of the techniques. The first consisted of 12 preference tests, in which all methods were compared. For each test, the subjects were asked to rank 5 different tracks, each of which contained a speech signal time-scaled using one of the methods A, B, C, D1 or D2. The second part consisted of eight pair comparisons, in which the proposed method (D) was compared to a traditional plain uniform-scaling method (A).

4.2 Test results

The results of the experiments show a clear preference for the proposed method, with 88% of listeners choosing a signal time-scaled by this method as their first choice in part one of the tests (Table 2).

Table 2. *First preference allocations*

Method:	**First Preferences**
D	88%
A or B or C	12%

The outcome of the overall rankings show a small improvement in quality of methods B and C compared to that of A, but methods D1 and D2 lead the field by a much more significant amount (Figure 3). This pattern is noticeable for all time-scaling factors investigated, as can be seen in Table 3.

Table 3. *Preference test results for different overall scaling factors*

	SF 2	**SF 2.5**	**SF 3**
First	D2	D2	D2
Second	D1	D1	D1
Third	A	C	C
Fourth	C	B	B
Fifth	B	A	A

Also evident from Table 3 is the deterioration in quality of method A as the time-scaling factor increases. This can be observed more clearly from the results of the second part of the test, in which 78% of listeners chose method D over method A (Table 4). The variation in this value with scaling factor forms the interesting result that, whereas method A decreases in quality as the scaling factor is increased, method D maintains a high quality output, as seen in Figure 4.

5. FUTURE WORK

Currently, various different techniques are being examined to determine the most efficient way of automatically detecting the different segmentals. Listening tests will be conducted to determine the optimum TSM factors to be used for each type of segmental. Real-time systems have been created within the research team, with implementation of the developed digital signal processing TSM algorithms, and a future task will be the development of a user-interface to enable the program to be used as a computer assisted speech therapy tool. This prototype can then be evaluated and tested by speech therapists and children with apraxia, and for this trial product, work is commencing on time-scaling of nursery-rhymes and the associated issues involved. Another future task will be to investigate the use of time-scale modification as an assistive-technology for other disabilities, such as stuttering, teaching pronunciation to the visually impaired, and in the rehabilitation of stroke victims.

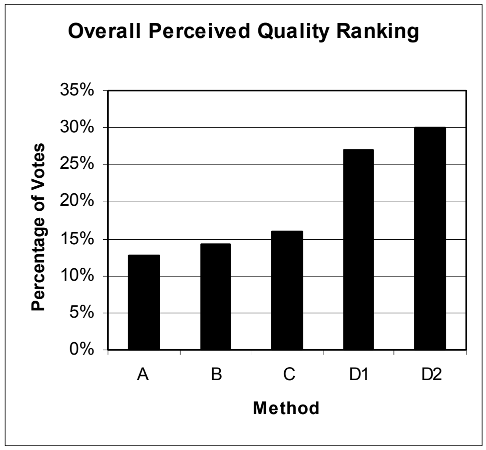


Figure 3. *Preference test results.*

Table 4. *Comparison preferences.*

Method:	**Preferences**
D	78%
A	22%

6. CONCLUSIONS

This paper discusses the merits of slowing down speech samples for use in a new computer-based speech therapy tool that allows poems and nursery rhymes to be slowed to any desired speed with minimal affect on the output quality. To be of benefit for this purpose the quality of the time-expanded speech needs to be very high. Several techniques to achieve high quality slowed speech were described and a new method using adaptive speech scaling was introduced. The tests carried out compared the quality of four different slow-down methods, namely uniform scaling, scaling only voiced segments, scaling only segments containing vowels and finally applying three different scaling factors to various types of speech segments while not scaling segments containing plosives.

The listening test results show that the proposed method using full segmental distinction is superior to the other methods and clearly delivers the best results. When large scaling factors are applied the advantage of the adaptive speech scaling method becomes even more apparent over traditional uniform speech scaling. The best results were achieved when the distance between the three different scaling factors increased exponentially. The proposed speech-adaptive slow-down system is therefore the most beneficial for the application in a CAST system.

Acknowledgements: This work was funded by the Enterprise Ireland administered Advanced Technologies Research Programme (ATRP) 2001, - Project ATRP/01/203, "DITCALL – Digital Interactive Tools for Computer Assisted Language Learning."

7. REFERENCES

T Ebihara, Y Ishikawa, Y Kisuki, T Sakamoto, T Hase (2000), Speech Synthesis Software with Variable Speaking Rate and its Implementation on a 32-bit Microprocessor, *19th IEEE International Conference on Consumer Electronics (ICCE 2000)*, Los Angeles Airport Marriott, USA

H Kuwabara (1997), Acoustic and Perceptual Properties of Phonemes in Continuous Speech as a Function of Speaking Rate, *Proc. Eurospeech 97*, pp. 1003-1006.

B Lawlor (1999), *Audio Time-Scale and Frequency-Scale Modification*, PhD Thesis, Department of Electrical Engineering, University College Dublin.

R Moir, M Sturm (2000), Time to Sing! CD, http://www.time2sing.

S Roucus, A M Wilgus (1985), High-Quality Time-Scale Modification for Speech, *IEEE Proceedings on Acoustics, Speech and Signal Processing,* pp. 493-496.

Consumer price data-glove for sign language recognition

T Kuroda[1], Y Tabata[2], A Goto[3], H Ikuta[4] and M Murakami

[1]Department of Medical Informatics, Kyoto University Hospital,
54, Kawahara-cho, Shogo-in, Sakyo-ku, Kyoto, 606-8507, JAPAN

[2]Kyoto College of Medical Technology,
1-3, Imakita, Koyama-higashi-machi, Sonobe-cho, Funai-gun, Kyoto, 62-0041, JAPAN

[3]AMITEQ Corp.,
5F, Hahshikan-LK-building, 1-6, Azuma-cho, Hachioji, Tokyo, 192-0082, JAPAN

[4]Teiken Limited,
1-6-7, Minami-honmachi, Chuo-ku, Osaka, 541-8587, JAPAN

[1]*Tomohiro.Kuroda@kuhp.kyoto-u.ac.jp,* [2]*yoshi-t@kyoto.medtech.ac.jp,*
[3]*amtq_goto@pop21.odn.ne.jp,* [4]*h.ikuta@teijin.co.jp*

[1]*www.kuhp.kyoto-u.ac.jp/~tkuroda,* [2]*www.kyoto.medtech.ac.jp*

ABSTRACT

A data-glove available for full degrees of freedom of a human hand is a key device to handle sign language on information systems. This paper presents an innovative intelligent data-glove named StrinGlove. StrinGlove obtains full degrees of freedom of human hand using 24 Inductcoders and 9 contact sensors, and encodes hand postures into posture codes on its own DSP. Additionally, the simple structure of the glove decreases the price. Several sign experts tried the prototype and the results show that the prototype has sufficient recognition rate as a sensor unit and sufficient comfortableness as a glove to wear.

1. INTRODUCTION

Sign linguistic engineering is a group of research to develop communication aid (called sign information system) for the Deaf and the Hearing Impaired, who have communication barrier in social lives, using information technology (Nagashima and Kanda 2001). Figure 1 shows the general view of sign information systems. Any sign information systems, such as virtual reality based sign language telephones, automatic translation system between sign languages and phonetic languages, and even sign language dictionary, share same basic structure. Most of sign information systems use three-dimensional computer graphic model of human beings called avatar to show signs. To obtain signs, though several systems using image recognition technologies, most of them are utilizing motion capture systems consists of data-globes and position and orientation sensors either magnetic-field-based, ultrasound-based, and image-based. The authors also developing sign language telephone named S-TEL (Kuroda et al, 1998) and consequent systems based on motion capture system consists of two data-gloves and three position and orientation sensors.

Thus, a data-glove available for full degrees of freedom of a human hand is a key device for sign information systems, and human computer interface for hearing impaired. However, data-gloves available on the current market are quite expensive, which cost as much as a luxurious car. Although several data-gloves are available in reasonable price, they can obtain only limited degrees of freedom. This poor variety of data-gloves prevents not only sign information systems but also other virtual reality based systems to spreading in the consumer market.

This paper presents an innovative data-glove named StrinGlove.

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

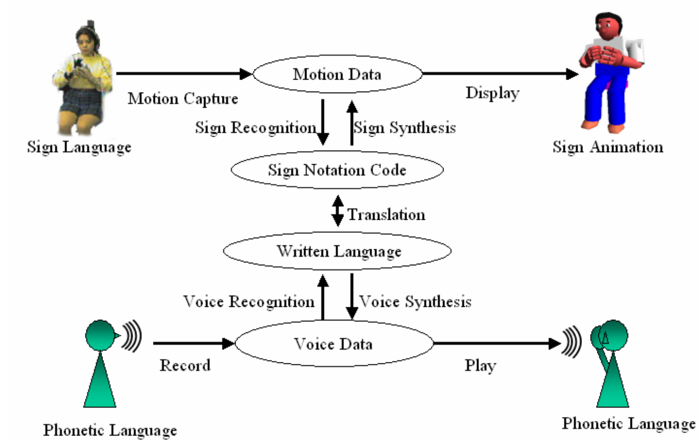


Figure 1. *General Overview of sign information systems*

2. FOREGOING DEVICES

In 1983, Jaron Lanier and Tom Jimmerman designed first glove type motion sensor to play music without touching any musical instruments, and 1984, they founded VPL Research to sell their sensor named DataGlove. DataGlove is a lightweight glove mounts ten fibre-optic cables along each finger. Each fibre-optic cable equips an LED at one end and phototransistor at the other end. When cable flexed, some of LED's light goes out from small scratches on surface of optic cable, so sensor obtains posture of a human hand. After invention of DataGlove, several types of data-gloves are presented (Zelter and Starman, 1994). Some gloves, such as Space Glove of Virtual Entertainment uses fibre-optic sensors as DataGlove, some Gloves, such as Nintendo's Powerglove and Altek's GLAD-IN-ART Glove utilizes piezoresistive sensors and some gloves, such as Exos' Dextrous Hand Master utilizes exoskeletal mechanical sensors. However, because of size and non-linearity of sensors, most of them have already gone away from market.

Most popular product on current market is CyberGlove of Immersion and 5DT Glove of 5th Dimension Technologies. These gloves utilize linear bend sensors mounted along each finger joints and present linear response. Although CyberGlove can obtain most of bending angles of human hand, it costs as much as a luxurious car. Nissho electronics' SuperGlove has same features as CyberGlove and cost much cheaper, it senses quite limited degrees of freedom and cannot applicable for motion capture on sign information systems.

3. DESIGN OVERVIEW OF STRINGLOVE

3.1 Sensing Bending Angles

When finger bends, wrinkles of back of finger joint stretches and distance along back of finger lengthen. Therefore, StrinGlove senses changes of distance between back of palm to each joint and fingertips as bends of each joint.

Figure 2 shows the basic structure of bending sensor of StrinGlove. The cable B consists of non-stretch fibre C and stretch fibre D connected by magnetizable material E. When joint A bends, the material E is pulled out from sensor tube. The motion of material E changes magnetic-flux density inside the sensor tube produced by active coil F, and the sensor obtains the change of magnetic-flux density as changes of inductance of coil F.

Figure 3 shows the detailed sensing mechanism of the sensor tube named Inductcoder® (Encoder Technology, 2001). Inductcoder induces magnetic field by giving single-phase alternating current $\sin(\omega t)$ on primary coil PW. Denoting displacement of material 15 as $0 \le \theta < 2\pi$, amplitudes of secondary coil S1~S4 become $\sin(\theta)$, $\cos(\theta)$, $-\sin(\theta)$, $-\cos(\theta)$ respectively. Therefore, S1 and S3, and S2 and S4 coupled as differential circuit as shown in Fig. 4, output alternating current A and B comes to $\sin(\theta)\cdot\sin(\omega t)$ and $\cos(\theta)\cdot\sin(\omega t)$ respectively. Thus, Inductcoder outputs two alternating currents as resolver. Therefore, Inductcoder can obtain absolute displacement of material 15 in high resolution using simple digital phase-shift detector.

As the Inductocoder gives linear output because of its simple mechanism, the glove just needs to obtain maximum and minimum sensing value to calibrate sensors. Therefore, StrinGlove obtains maximum and minimum value during certain period of time after its initialization and calibrate itself automatically. Therefore, users has no needs to perform any special calibration process before using the sensor instead of opening and closing one's hands for several times.

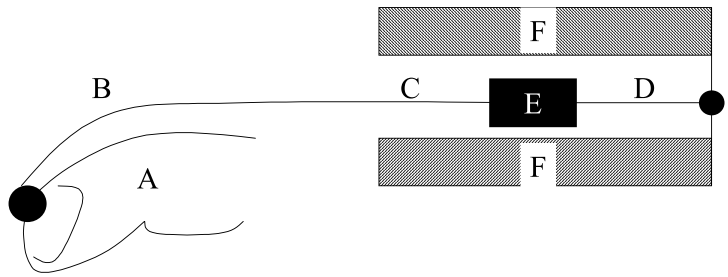


Figure 2. *Basic structure of bending sensor.*

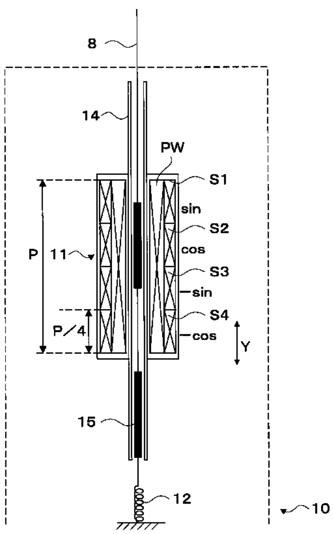


Figure 3. *Detailed design of Inductcoder.*

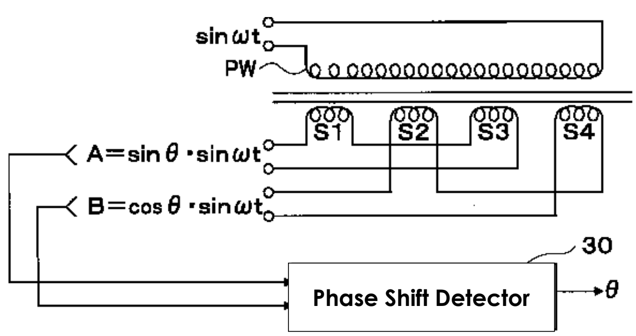


Figure 4. *Circuit diagram of Inductcoder.*

StrinGlove equips 24 Inductcoders to obtain 22 degrees of freedom of human hand as shown in Fig. 5. As wrist bends both inner and outer directions, four sensors are equipped to obtain two degrees of freedom of wrist. As the external diameter of sensor tube is only three mille-meters, most of sensors can be equipped on back of palm without disturbing any hand motion. Depends on maximum changes of distance between finger tip and perm, the standard length of sensor tube is 44.5 mille-meters although length of some sensor tubes to obtain bends of metacarpophalangeal joint (MPJ) are 30.5 mille-meters. The Inductcoder fasten its cable with small clasp as shown in Fig. 6 in both ends, users can adjust or replace sensor cable in case it cut.

3.2 Sensing contact between fingertips

To obtain contacts between fingertips using bending sensors are quite difficult as size of human hands varies widely. Therefore, the authors decided to put contact sensors on fingertips. PinchGlove of Fakespace or TouchGlove of Infusion Systems obtains pressures given by sensors or contacts of fingers. However, human beings not always really touch fingertips firmly even when one intends to do so. Therefore, StrinGlove utilizes approaching sensor using magnetic coils.

The authors examined hand postures in sign languages around the world and found that nine contact sensors as shown in Fig. 7 is sufficient to obtain any hand posture. To give different frequency on each active sensor coils on fingers, StrinGlove may distinguish to which finger the thumb touches from sensor value from reactive sensor coil mounted on tip of thumb. Additionally, frequency shift of each active sensor coils may obtain crossings of fingers such as posture shown in Fig.8, indicates Japanese character "ra".

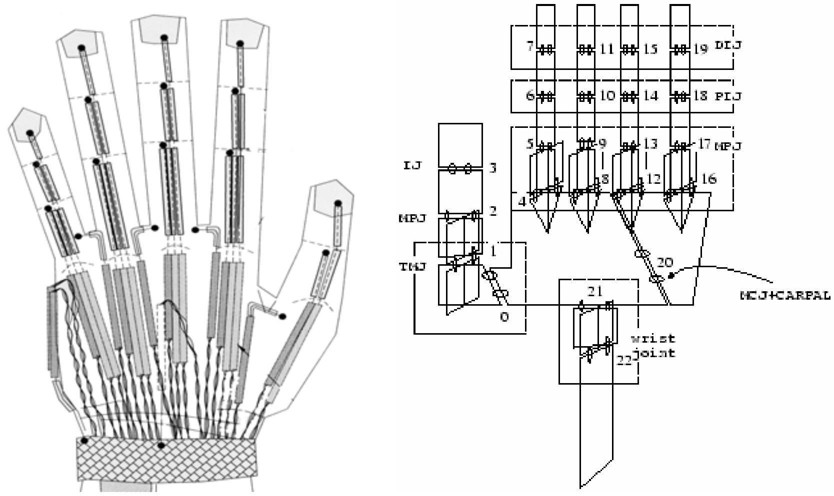


Figure 5. *Bend sensor arrangement and target hand model.*

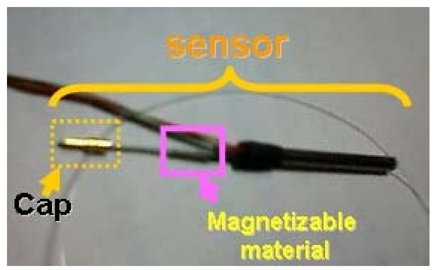


Figure 6. *Clasps to fasten sensor cable.*

3.3 Fabrics

Make glove type sensor sensitive enough, glove must fit perfectly on human hand. On the other hand, glove mustn't disturb free motion of human hand. Therefore, soft and stretchable fabric should be used for glove type sensor. The authors realized alpha version glove using soft fabrics. However, trial performed by sign experts reveals that too soft fabric make the glove difficult to wear on and off. Because the glove made of too soft fabric cannot hold its shape at any moment, and cannot hold sensors mounted on the glove. Moreover, sensor cables are cut out from sensor tube easily because the glove itself stretches too much when user wear off the glove.

To overcome these problems, the authors used two different fabrics; semi-stretching fabric, which stretches only for a single direction, and stretching fabric, which stretches in any directions. The glove uses stretching fabric for backside and semi-stretching fabric for palm side to avoid stretching along finger direction. Thus, although the glove has enough elasticity to fit user's hand and sense stretches of sensor cable, any sensor cables won't be cut.

Additionally, StrinGlove mounts sensor units by Velcro in order to enable to take all the sensors away and to wash the sensor glove. As no sensor gloves in market is washable, StrinGlove has high advantages for public use.

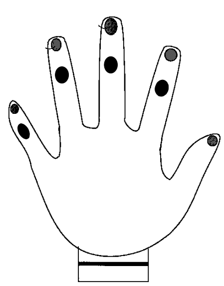


Figure 7. *Arrangement of contact sensors.*

Figure 8. *Japanese character "ra" (Yohei).*

3.4 Encoder

Sign language is a group of most sophisticated coding system of human motion. The authors reviewed existing notation system for sign languages, such as HamNoSys (Hanke 2004), sINDEX (Kanda 1997), and so on, and defined hand shape coding system, which covers most of intentional hand motion. Table 1 shows a part of developed coding system. As the coding system is rather simple, StrinGlove encode hand posture into notation code using DSP mounted on the glove itself using vector quantization.

This automatic encoding function decrease calculation load of central CPU, so it eases to develop sign information system and any virtual reality applications. StrinGlove, as matter of course, presents sensor data as it is in twelve bits depending on given order. Figure 9 shows the beta prototype.

Table 1. *Example codes of developed coding system.*

Code	Finger states	Code	Finger states
H	Full stretches of all fingers	**A**	Abduction between fingers
B	Full bending of all fingers	**G**	Abduction of first joint in thumb, and bending of second, third joints in thumb
b	Bending of all fingers	**I**	Adduction of first joint in thumb, and stretch of second, third joints in thumb
F	Bending of first and second joints	**T**	Touch between fingertips

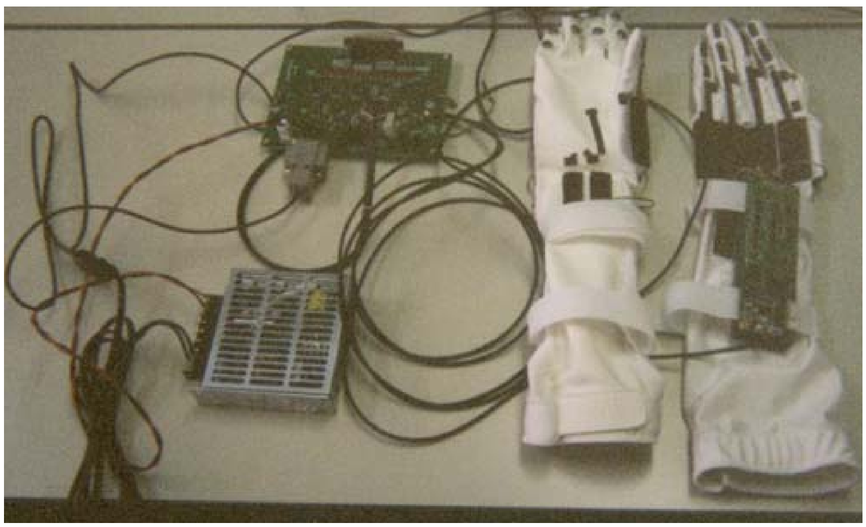


Figure 9. *Prototype of StrinGlove.*

4. EVALUATION

Ten sign experts including four deaf evaluated the beta prototype. All subjects were women. Figure 10 shows several snaps from the trial.

The subjects compared comfortableness of the prototype in comparison with CyberGlove. The subjects were asked to wear both gloves and to express several words. After the trial, the subjects were interviewed. All subjects claimed that they could play with the prototype without any physical fatigue although they cannot play with the CyberGlove for a long time, because the prototype allows subjects to bend their finger without hindrance. This claim ensures the comfortableness of the prototype.

The accuracy of coding function of the prototype is examined. The subjects are asked to express 48 finger characters with the prototype on, and the sensor outputs are compared with correct notation code of each finger characters. The accuracy of encoding was about 85%. Although two pairs of prototypes in different size are prepared, pinkies of several subjects didn't fit with the prototype. Thus, ill-fitting glove seems lower the recognition rate. Although ill-fitting glove lowered the accuracy, the result shows that the prototype encodes human hand posture in sufficient accuracy.

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

Additionally, the prototype obtains finger character “ra” (see Fig. 8), which conventional data-gloves cannot distinguish, without any error throughout the trial. This result clears that the prototype utilizing bend and contact sensors to obtain hand posture has clear advantage in comparison with conventional data-gloves.

Thus, the experimental results show that StrinGlove, which is comfortable to wear and has sufficient coding function, has clear advantage in comparison with conventional data-gloves.

Figure 10. *Snaps from evaluation of the prototype.*

5. CONCLUSION

This paper presents an innovative intelligent data-glove named StrinGlove. StrinGlove captures motion of human hand using 24 bending sensors and nine contact sensors and encode given postures by itself. Thus, StrinGlove eases to develop sign information system and any virtual reality system. Additionally, the simple structure of StrinGlove allows cutting the cost of the glove into reasonable price and makes itself washable. Therefore, the authors believe that StrinGlove accelerate development of commercial sign information systems and virtual reality systems.

Acknowledgements: The authors appreciate for warm support of Fujita Corp. (JP), Digital Research Inc. (JP), Data Processing and Research Inc. (VN), TMA Solutions (VN), Polygon Pictures (JP), Matsumura Corp. (JP), and NIF Ventures Co. Ltd. (JP). The authors also thank “FUREAI”, sign language learning club of Sakai city. This research is partly funded by Japan Science and Technology Agency. All patents about Inductcoder® and StrinGlove are pending.

6. REFERENCES

Encoder Technology (2001) A Discussion of Encoder Technology’s Converter and Impedance Detector, http://www.encoder-technology.com/products/amitech/inductcoderwhitepaper.pdf

D Zeltzer and D J Sturman (1994), A survey of glove-based input, *IEEE Computer Graphics and Applications*, **14**, *1*, pp.30-39

K Kanda, Y Nagashima, A Ichikawa, and M Terauchi (1997), A proposal of Standard Labelling System for JSL: Signdex V.1, *Proc. HCI 97*, p.15

T Hanke (2004) HamNoSys &hibar; representing sign language data in language resources and language processing contexts, *Proc. LREC 2004*, pp.1-6

T Kuroda, K Sato and K Chihara (1998), S-TEL an avatar based sign language telecommunication system, *Int, J. Virtual Reality*, **3**, *4*, pp. 21-27

Y Nagashima and K Kanda (2001), Present state and issue of sign linguistic engineering, *Proc. HCI International*, **3**, pp. 387-391

Yohei, Yubimoji –BIG GATE–, http://www2.incl.ne.jp/~yohei/

Synthesis of virtual reality animation from sign language notation using MPEG-4 body animation parameters

M Papadogiorgaki[1], N Grammalidis[2], N Sarris[3] and M G Strintzis[4]

[1,2,4]Informatics and Telematics Institute
1st Km Thermi-Panorama Road, 57001 Thermi-Thessaloniki, GREECE

[3]Olympic Games Organizing Committee, Athens 2004
Iolkou 8 & Filikis Etairias, 14234 Nea Ionia, GREECE

[1]mpapad@iti.gr, [2]ngramm@iti.gr, [3]NESarris@athens2004.com, [4]strintzi@iti.gr

[1,2,4]www.iti.gr, [3]www.athens2004.gr

ABSTRACT

This paper presents a novel approach for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Face and Body Animation. Sign Language notation, in the well-known Sign Writing system, is provided as input and is initially converted to SWML (Sign Writing Markup Language), an XML-based format that has recently been developed for the storage, indexing and processing of Sign Writing notation. Each basic sign, namely sign box, is then converted to a sequence of Body Animation Parameters (BAPs) of the MPEG-4 standard, corresponding to the represented gesture. In addition, if a sign contains facial expressions, these are converted to a sequence of MPEG-4 Facial Animation Parameters (FAPs), while exact synchronization between facial and body movements is guaranteed. These sequences, which can also be coded and/or reproduced by MPEG-4 BAP and FAP players, are then used to animate H-anim compliant VRML avatars, reproducing the exact gestures represented in the sign language notation. Envisaged applications include interactive information systems for the persons with hearing disabilities (Web, E-mail, info–kiosks) and automatic translation of written texts to sign language (e.g. for TV newscasts).

1. INTRODUCTION

The SignWriting system is a writing system for deaf sign languages developed by Valerie Sutton for the Center of Sutton Movement Writing, in 1974 (SignWriting site, 2004). A basic design concept for this system was to represent movements as they are visually perceived, and not for the eventual meaning that these movements convey. In contrast, most of the other systems that have been proposed for writing deaf sign languages, such as HamNoSys (the Hamburg Notation System) or the Stokoe system employ alphanumeric characters, which represent the linguistic aspects of signs. Almost all international sign languages, including the American Sign Language (ASL) and the Brazilian Sign Language (LIBRAS), can be represented in the SignWriting system. Each sign-box (basic sign) consists of a set of graphical and schematic symbols that are highly intuitive (e.g. denoting specific head, hand or body postures, movements or even facial expressions). The rules for combining symbols are also simple, thus this system provides a simple and effective way for common people with hearing disabilities that have no special training in sign language linguistics, to write in sign languages. Examples of SignWriting symbols are illustrated in Figure 1.

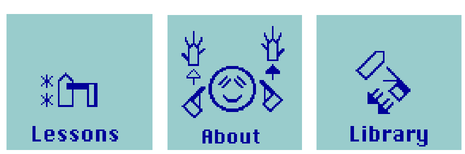


Figure 1. *Three examples of representations of American Sign Language in SignWriting system.*

An efficient representation of these graphical symbols in a computer system should facilitate tasks as storage, processing and even indexing of sign language notation. For this purpose, the SignWriting Markup Language (SWML), an XML-based format, has recently been proposed (Costa, 2001). An online converter is currently available, allowing the conversion of sign-boxes in SignWriting format (produced by SignWriter, a popular SignWriting editor) to SWML format.

Another important problem, which is the main focus of this paper, is the visualization of the actual gestures and body movements that correspond to the sign language notation. Grieve-Smith (2001) presented a thorough review of state-of-the art techniques for performing synthetic animation of deaf signing gestures. Traditionally, dictionaries of sign language notation contain videos (or images) describing each sign-box, however the production of these videos is a tedious procedure and has significant storage requirements. On the other hand, recent developments in computer graphics and virtual reality, such as the new Humanoid Animation (H-Anim) (H-anim, 2004) and MPEG-4 SNHC (MPEG-4 document, 1999) standards, allow the fast conversion of sign language notation to Virtual Reality animation sequences, which can be easily visualized using any VRML-enabled Web browser.

In this paper, we present the design, implementation details and preliminary results of a system for performing such a visualization of sign-boxes, available in SWML. The proposed technique first converts all individual symbols found in each sign box to sequences of MPEG-4 Face and Body Animation Parameters. The resulting sequences can be used to animate any H-anim-compliant VRML avatar using MPEG-4 SNHC BAP and FAP players, provided by EPFL. The system is able to convert all hand symbols as well as the associated movement, contact and movement dynamics symbols contained in any ASL sign-box. Manual (hand) gestures and facial animations are currently supported, while we plan to implement other body movements (e.g. torso) in the near future. The proposed technique has significant advantages:

- Web- (and Internet-) friendly visualization of signs. No special software has to be installed except a VRML plug-in to a Web browser,
- Allows almost real-time visualization of sign language notation, thus enabling interactive applications,
- Avatars can easily be included in any virtual environment created using VRML, which is useful for a number of envisaged applications, such as TV newscasts, automatic translation systems for the deaf, etc.
- Efficient storage and communication of animation sequences, using MPEG-4 coding techniques for BAP/FAP sequences.

Significant similar work for producing VRML animations from signs represented in the HamNoSys transcription system to VRML has been carried out by the EC IST ViSiCAST project (Kennaway, 2001), and its follow-up project “E-Sign”(E-sign site, 2004). Current extensions of HamNoSys are able to transcribe all possible body postures, movements and facial expressions (Hanke, 2002) and significant work towards supporting MPEG-4 BAPs has been made. The main contribution of the proposed approach in this paper is the attempt to work towards the same direction for the most common and popular representation of Sign Languages, which is the SignWriting notation system.

The paper is organized as follows: In Section 2, the proposed technique for converting SWML sign boxes to MPEG-4 Face and Body Animation Parameters is described. The synthesis of animations for H-anim avatars and the design of the experimental “Vsigns” Web page to evaluate the sign synthesis results are outlined in Section 3, while discussion and future work is presented in Section 4.

2. CONVERSION OF SWML SIGN BOXES TO MPEG-4 FACE AND BODY ANIMATION PARAMETERS

In this Section, we briefly describe the procedure to convert SignWriting notation in SWML format to MPEG-4 Face and Body Animation Parameters. SWML (SWML Site, 2004) is an XML-based format for the representation of SignWriting notation described by the SWML DTD (currently version 1.0 draft 2) (Costa, 2001).

Each SWML *signbox* consists of a set of symbols, which is specified using the following fields:

a) A shape number (integer) specifying the shape of the symbol,
b) A variation parameter (0 or 1 for hand symbols / 1,2 or 3 for movement and punctuation symbols) specifying possible variations (complementary transformations) of the symbol,

c) A fill parameter (0,1,2 or 3 for hand and punctuation symbols / 0,1 or 2 for movement symbols) specifying the way the shape is filled, generally indicating its facing to the signer,
d) A rotation parameter (0-7) specifying a counter-clockwise rotation applied to symbol, in steps of 45 degrees,
e) A transformation flip parameter (0 or 1) indicating whether the symbol is vertically mirrored or not, relatively to the basic symbol and, finally,
f) The x and y coordinates of the symbol within the sign box.

For sign synthesis, the input for the sign synthesis system consists of the SWML entries of the sign boxes to be visualized. For each sign box, the associated information corresponding to its symbols is parsed.

Currently, symbols from the 1995 version of the Sign Symbol Sequence (SSS-1995) are supported. This sequence comprises an "alphabet" of the SignWriting notation system, while true images (in gif format) of each symbol contained in this sequence are available in (SWML site, 2004). The proposed system is able to convert

- All 106 hand symbols,
- All 95 (hand) movement symbols
- Two punctuation symbols (180,181), which contain synchronization information.
- 27 Facial expression/animation symbols

Other punctuation symbols as well as symbols that represent torso and shoulder movements (12 symbols) are currently not implemented (decoded) by the system. Information from symbols, within each sign-box, that are supported by the sign synthesis application, i.e. hand symbols as well as corresponding movement, contact and movement dynamics symbols, is then used to calculate the MPEG-4 Face and Body Animation Parameters.

The issue body modelling and animation has been addressed by the Synthetic/Natural Hybrid Coding (SNHC) subgroup of the MPEG-4 standardization group (MPEG-4 document, 1999). More specifically, 168 Body Animation Parameters (BAPs) are defined by MPEG-4 SNHC to describe almost any possible body posture. In addition, 68 Face Animation Parameters (FAPs) are used to describe almost any possible facial expression. Most BAPs denote angles of rotation around body joints, while FAPs usually denote movements of specific facial features (Facial Definition Points, FDPs) along a pre-determined axis in 3-D space.

The conversion of the symbols contained in a SWML sign box to BAP sequences starts by first examining the symbols contained within the input sign box. If no symbols describing dynamic information such as hand movements, contact or synchronization exist, the resulting BAP sequence corresponds to just one frame (i.e. a static gesture is reproduced). Information provided by the fields of the (one or two) hand symbols, contained in the sign box, is used to specify the BAPs of the shoulder, arm, wrist and finger joints. On the other hand, if symbols describing dynamic information exist, the resulting BAP sequence contains multiple frames, describing animation key-frames (i.e. a dynamic gesture is reproduced). The first key-frame is generated by decoding the existing hand symbols, as in the case of static gestures. Since the frame rate is constant and explicitly specified within a BAP file, the number of resulting frames may vary, depending on the complexity of the described movement and its dynamics. Synchronization symbols and contact also affect the represented movement and in some cases require special treatment.

When a signbox contains facial expression or animation symbols, the corresponding FAP frame(s) are determined by predefined lookup tables, which provide the FAP values defining one or more FAP frames per facial animation symbol. When two or more facial expression symbols co-exist within the same sign-box, these may either define an animation sequence or have to be combined all together (if each symbol activates different FAPs). The latter case, which is more common, is currently supported by the proposed system.

Smooth and natural-looking transitions between the Face and Body Animation parameters corresponding to each signbox is achieved by generating additional intermediate frames using a FAP/BAP interpolation procedure. A linear interpolation function is used to generate additional FAP/BAP frames to implement:

a) The transition between the neutral face/body position and the first frame of the first sign-box
b) The transition between the end frame of one signbox and the start frame of the next signbox
c) The transition between the end frame of the last sign-box and the neutral body position.

Furthermore, in order to achieve Face/Body synchronization:

a) The frame rates defined for the FAP and BAP sequences should be equal

b) The number of generated FAP frames generated for each sign-box should be always equal to the corresponding number of BAP frames. In order to achieve this goal, the BAP frame sequence is first generated and then specific linear interpolation procedures are used to generate the FAP frame sequence.

A block diagram of the proposed system for processing each sign-box is illustrated in Figure 2, while additional details about the generation of BAPs for static and dynamic gestures as well as the generation of FAPs for gestures containing facial expressions/animations are provided in the following Subsections.

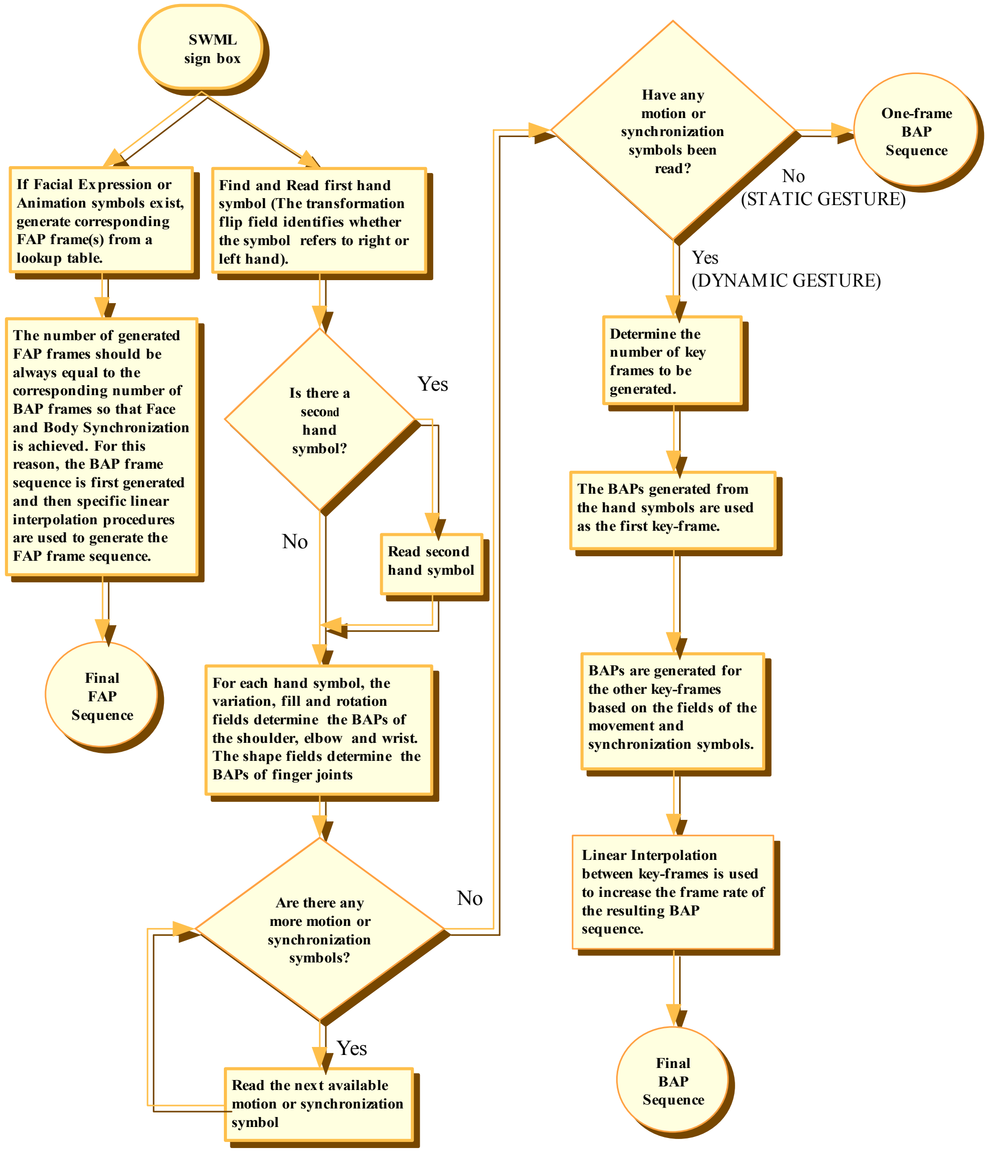


Figure 2. *A block diagram of the proposed system.*

2.1 Static gestures

In the following, the procedure to extract useful information from the SWML representation of a hand symbol is summarized:

Initially, the binary "transformation flip" parameter is used to identify whether the symbol corresponds to the left or right hand. Then the fill and variation parameters of each symbol are used to determine the animation parameters of the shoulder and elbow joints:

- If (variation, fill)=(0,0),(0,1) or (1,3) then the axis of the arm is parallel to the floor (floor plane).

Proc. 5th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, UK, 2004

- If (variation, fill)=(1,0),(1,1) or (1,2) then the axis of the arm is parallel to the human body (wall plane)
- If (variation, fill)=(1,0) or (1,3) then the signer sees his palm
- If (variation, fill)=(1,1) or (0,0) then the signer sees the side of his palm
- If (variation, fill)=(1,2) or (0,1) then the signer sees the back of his palm

In addition, the rotation parameter is used to determine the animation parameters of the wrist joint:

- If the signer sees the side of his palm, the rotation value (multiplied by 45 degrees) is used to define the R_WRIST_FLEXION BAP (for the right hand) or the L_WRIST_FLEXION BAP (for the left hand).
- In the other two cases (signer sees his palm or the back of his palm), the rotation value (multiplied by 45 degrees) is used to define the R_WRIST_PIVOT BAP (for the right hand) or the L_WRIST_PIVOT BAP (for the left hand).

Finally, the symbol shape number is used to specify the animation parameters corresponding to finger joints, using look-up tables of BAP values corresponding to each symbol.

If the sign box contains a second hand symbol, similar procedures are used to extract the Body Animation Parameters of the other hand. After the processing of all existing hand symbols, all Body Animation Parameters corresponding to shoulder, elbow, wrist and finger joints are determined and stored.

2.2 Dynamic gestures

MPEG-4 standard allows the description of human body movement using a specific set of Body Animation Parameters corresponding to each time instant. Systems like SignWriting that use a high level animation description define movement by specifying a starting and an ending position, in case of simple motion with constant velocity, or the full trajectory, in case of more complex motion. However, the description of complex motion is also possible by specifying a number of intermediate key-frames. In the following, the procedures for generating these BAP key-frames are briefly described.

When all movement description symbols have been identified, the shape number field identifies their shapes (i.e. the type of movement). First, the total number of key-frames to be produced is specified, based on the number and nature of the available movement, movement dynamics, contact, and synchronization symbols. More specifically, a look-up table is used to define an initial number k of key frames for each movement symbol. Furthermore, the fill parameter specifies whether the motion is slow, normal or fast. In addition, some symbols explicitly specify the movement duration. For this reason, a classification of such symbols into three categories has been defined and a different duration value D is defined for each category:

- Slow motion (D=3)
- Normal motion (D=2)
- Fast motion (D=1)

The total number of frames to be generated when only one motion symbol exists is $N=kDP$, where P is a fixed multiplier (e.g. P=10). If the number of such symbols is more than one, the total number of key-frames is the maximum between the numbers of key-frames, corresponding to each symbol. Finally, if the sign box contains a contact symbol, the total number of frames is increased by two (in case of simple contact) or four (in case of double contact).

The initial key-frame is generated by decoding the available hand symbols, exactly as in the case of static gestures. The rotation and transformation flip fields specify the exact direction of movement. Also, the variation field specifies whether the right or the left hand performs the movement. Using information from all available movement, contact and synchronization symbols, the other BAP key-frames of the specific dynamic gesture are then generated using a specific function for each key-frame. Synchronization (Movement Dynamics) symbols (180,181 and 182) are handled in a similar way as movement symbols but an exception exists for the "Un-even alternating" symbol, where first one hand moves, while the other hand is still and then the opposite. To handle this case the total number of key frames is doubled ($N=2kDP$). To produce the first kDP frames, BAPs are generated only for the first hand, so the second hand remains still. Then, BAPs are generated only for the second hand, to produce the next kDP frames, so the first hand remains still.

Finally, when the BAPs for all key-frames have been computed, BAP interpolation is used to increase the frame rate of the resulting BAP sequence. This interpolation procedure results to smoother transitions between key frames.

Interpolation is generally achieved by approximating the motion equation using a mathematical function and then re-sampling this function to obtain the desired intermediate positions at intermediate time instants. Various interpolation functions can be selected in order to improve results. Since Body Animation Parameters represent rotations around specific joints, quaternion interpolation was seen to provide good results [8], but the complexity of the method is increased. For this reason, a linear interpolation technique was applied, which was seen to be very efficient for most signs, since key-frames have been selected so as to simplify the movement description between consecutive key-frames.

2.3 Gestures containing facial expressions-animations.

The generation of the FAP frame sequence is performed after the generation of the BAP frame sequence, so that the total number of generated FAP frames is exactly the same as the total number of BAP frames. For each sign-box, the FAP key-frames are determined, based on the existing facial expression/animation symbols, from predefined lookup tables for each symbol. The number of FAP key-frames, $N_{FAP\_keyframes}$, is generally much smaller than the total number of BAP frames N_{BAP} that have been already generated using the procedures described in the previous Subsections. Therefore, if $FAP(k)$, $k = 0, ..., (N_{BAP} - 1)$ denotes the vector of FAPs corresponding to frame k, the FAP key frames are first positioned every $step = N_{BAP}/(N_{FAP\_keyframes} - 1)$ frames:

$$FAP(i*step) = FAP\_keyframe(i*step),\ i = 0, ..., /(N_{FAP\_keyframes} - 1)$$

Then, each of the remaining FAP frames is determined using linear interpolation between the two closest available FAP key frames.

3. SYNTHESIS OF ANIMATIONS USING H-ANIM AVATARS

The "EPFLBody" BAP player (Vergnenegre, 1999), developed by the École Polytechnique Fédérale Lausanne (EPFL) for the Synthetic and Natural Hybrid Coding (SNHC) subgroup of MPEG-4 was used to animate H-anim-compliant avatars using the generated BAP sequences. Since most BAPs represent rotations of body parts around specific body joints, this software calculates and outputs these rotation parameters as animation key-frames to produce a VRML ("animation description") file that can be used for animating any H-anim-compliant VRML avatar. The "Miraface" FAP player, also developed for MPEG-4 SNHC, by MIRALab, University of Geneva and LIG, EPFL was used for Facial Animation. This software had to be modified so that:

a) VRML animation output is produced using one CoordinateInterpolator node per face model vertex. A problem with the chosen implementation is that the computational demands for the hardware that is reproducing these animations are increased. A possible solution for this problem that should be investigated in the future is to add CoordinateInterpolator nodes only for the points that have actually been moved.
b) The face model to be animated using the FAP frame sequence was attached to the body to be animated using the BAP frame sequence. Some slight modifications of the VRML face model were also required (e.g. addition of teeth).

Two frames from resulting animations are illustrated in Figure 7.

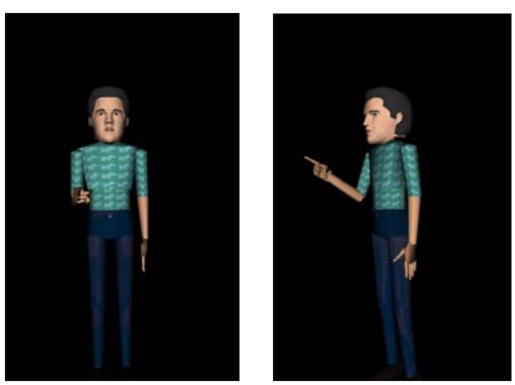


Figure 7. *Animation of the "You" sign in ASL using an H-anim avatar*

By including a VRML TouchSensor Node within the VRML file describing the H-anim avatar, the viewer can interactively start and/or replay the animation sequence, by clicking on the avatar. The viewer can also interact by zooming in and out to any specific body region and/or by rotating and translating the model within the 3-D space, in order to fully understand the represented sign.

Furthermore, further evaluation of the proposed sign synthesis system was possible by developing an online system (Vsigns site, 2004) for converting text to Sign Language notation and corresponding VRML animation sequences for H-anim compliant avatars. The application, whose interface is illustrated in Figure 8, is currently based on a 3200-word SWML dictionary file, obtained by the SWML site, which has been parsed and inserted into a relational database. The user is allowed to enter one or more words, which are looked up in this dictionary. If more than one entry is found, all possible interpretations are presented to the user, so that he can choose the desired one. On the other hand, if no entries are found for a specific word, the word is decomposed using its letters (finger-spelling). In any case, the user may choose whether to include a particular term to the selected terms to be used for sign synthesis or not. The user then selects a column corresponding to an H-anim compliant avatar, which is used for sign synthesis of the selected term or terms. A fourth column ("Baxter FBA") allows the user to observe facial animation in addition to body animation, using the modified "Baxter avatar". Furthermore, the user may produce and display the corresponding sign(s) in SignWriting format (in PNG format) and SWML for a specific term or the selected terms.

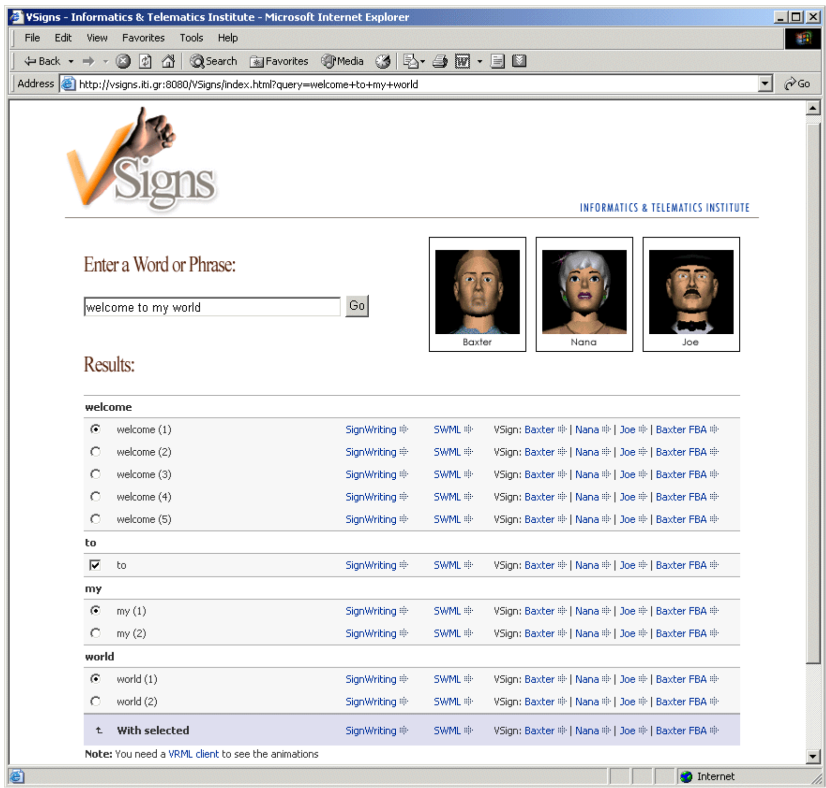


Figure 8. *Example query: "Welcome to my world". The user may then select the desired terms and then produce and display sign synthesis results using the selected words or the entire phrase, using any of the available H-anim avatars.*

This experimental Web application has already allowed us to identify significant problems with the synthesis of static and dynamic gestures, which have to be solved in the future, e.g. when contacts and complex movements are involved. A major problem that has to be solved occurs when the sign-box contains contact symbols. In that case the touch between the hands, or the hand and the face is difficult to be achieved. Problems may also occur for complex movements, when the inclinations of the hand joints, which have been estimated in each key frame, are not accurate enough for the exact description of the movement. Both problems can be solved in the future by using inverse kinematics methods, as in (A B. Grieve-Smith, 2001).

Further evaluation is planned for the future, using Greek and International SignWriting users, and attempts will be made to solve possible problems in the reproduction of specific signs. Although these problems indicate that much more work is needed for correct synthesis of all signs, we believe that with this Web tool, a very important step towards automatic Text to Sign synthesis has been made.

4. DISCUSSION AND FUTURE WORK

A novel approach for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Body Animation has been presented. The system is able to convert almost all hand symbols as well as the associated movement, contact and movement dynamics symbols contained in any ASL sign-box. Furthermore, most facial expression and animation symbols are also supported, while torso movements will be also supported in the near future. Some facial expressions, e.g. cheek wrinkles, have not been implemented, since no FAPs exist to produce such movements.

Results are satisfactory and are currently being evaluated by SignWriting users and experts, so that problems associated with specific SignWriting symbols are identified and solved. In the future, improved reproduction of difficult movements (e.g. touching) will be made possible using inverse kinematics techniques.

A short-term goal is to design other practical applications of the proposed system, either as a “plug-in” to existing applications (e.g. sign language dictionaries) or as a stand-alone tool for creating animations for TV newscasts (e.g. weather reports). Particular emphasis will be given in applications that can be used and evaluated by the Greek Sign Language community, thus a dictionary of Greek Sign language, in SignWriter notation, is planned to be supported in the near future.

Acknowledgements: This work was supported by FP6 IST Network of Excellence “SIMILAR” -”The European taskforce creating human-machine interfaces SIMILAR to human-human communication”. The authors would also like to thank Lambros Makris for designing and developing the “Vsigns” Web Page.

5. REFERENCES

Official Sign Writing site, http://www.signwriting.org/

“Generic coding of audio-visual objects - Part 2: Visual”, *MPEG Document ISO/IEC JTC 1/SC 29/WG11* N3056, Maui, December 1999

Official SignWriting site, http://www.signwriting.org/

Official Site of SWML, http://swml.ucpel.tche.br/

Moving Pictures Experts Group, (1999). Generic coding of audio-visual objects - Part 2: Visual, MPEG Document ISO/IEC JTC1/SC29/WG11 N3056. Maui.

F Vergnenegre, Tolga K. Capin, and D. Thalmann (1999). Collaborative virtual environments-contributions to MPEG-4 SNHC. ISO/IEC JTC1/SC29/WG11 N2802, http://coven.lancs.ac.uk/mpeg4/

A B. Grieve-Smith (2001). SignSynth: A Sign Language Synthesis Application Using Web3D and Perl. Gesture Workshop, London, pp. 134-145

R. Kennaway (2001). Synthetic Animation of Deaf Signing Gestures. Gesture Workshop, pp. 146-157, London.

Antonio Carlos da Rocha Costa, Cracaliz Pereira Dimuro (2001). Supporting Deaf Sign Languages in Written Form on the Web. The SignWriting Journal, Number 0, Article 1, July. http://gmc.ucpel.tche.br:8081/sw-journal/number0/article1/index.htm.

M. Preda, F. Preteux (2001). Advanced virtual humanoid animation framework based on the MPEG-4 SNHC standard. Proceedings EUROIMAGE International Conference on Augmented, Virtual Environments and Three-Dimensional Imaging (ICAV3D’01), Mykonos, Greece, pp. 311-314.

Humanoid Animation Standard Group. Specification for a Standard Humanoid: H-Anim 1.1. http://h-anim.org/Specifications/H-Anim1.1/

Official site of E-sign (Essential Sign Language Information on Government Networks) project. http://www.visicast.sys.uea.ac.uk/eSIGN/

Th. Hanke (2002). iLex - A tool for sign language lexicography and corpus analysis. In: Proceedings of the Third International Conference on Language Resources and Evaluation, Las Palmas de Gran Canaria, Spain., pp. 923–926.

Vsigns page, http://vsigns.iti.gr

ICDVRAT 2004

Session IX. Interfacing to and Navigation within Virtual Environments

Chair: Tony Lewis- Brooks

Using switch controlled software with people with profound disabilities

N Anderton[1], P J Standen[2] and K Avory[3]

[1,2,3]Division of Rehabilitation and Ageing, University of Nottingham,
ADRU, B Floor, QMC, Clifton Boulevard, Nottingham NG7 2UH, UK

n.anderton@nottingham.ac.uk, p.standen@nottingham.ac.uk, Karenavory@yahoo.com

[1,2,3]www.nottingham.ac.uk/rehab

ABSTRACT

Micro switches used to control sound and light displays increase the level of activity of people with profound intellectual disabilities and provide a means by which they can exert some control over their environments. This study set out to i) explore whether people with profound disabilities could learn to use a simple game controlled by a single micro switch and displayed on a normal computer monitor; ii) document what activities on the part of a tutor best facilitated the performance of the learner. Four men and three women aged between 24 and 46 years with profound disabilities completed eight twice weekly sessions when they were given the opportunity to play a computer game that could be operated by a large jelly bean switch. A tutor sat next to them throughout the sessions and each session was recorded on videotape. Tapes were analysed for the help given by the tutor, use of the switch and duration of attention. Although the game was too difficult for them, all participants increased the percentage of time during the session in which they looked at the monitor and for all of them there were at least two sessions when their switch pressing became a consequence of the tutor's activity.

1. INTRODUCTION

People with profound intellectual disabilities are unable to cope independently and their level of ability together with the presence of additional disabilities makes it difficult for them to be involved in conventional activities (Lancioni, et al, 2002). This results in their isolation, passivity and failure to obtain desirable environmental stimulation. In their natural environments they receive very low levels of interaction with others, typical exchanges lasting less than 60 seconds (Clegg, Standen & Cromby, 1991). However, more social approaches have been recorded when they are involved in some form of training (Landesman, 1987).

Clegg, Standen & Cromby (1991) demonstrated the importance of social routines in increasing positive behaviour in people with profound disabilities. They investigated the effect of five different staff strategies on behaviour and found that increased positive behaviour was associated with staff talking with them and social routines. Social routines are exchanges between individuals such as tickling games or other turn taking rituals. However, staff are often unsure of how to talk with someone who does not use verbal or symbolic communication or how to initiate social routines especially with clients with whom they are unfamiliar. This may explain why the majority of the interactions they have with those for whom they are responsible centre on physical care giving. The research by Landesman (1987) suggests that a learning task could provide a focus for developing the social routines valuable for the well-being of those with profound disabilities.

There is now considerable evidence supporting the use of computer technology for people with intellectual disabilities (Standen, Brown & Cromby, 2001) especially as an aid to learning. There has consequently been a growth in the design of multimedia tailored to their needs (e.g. Brown, Shopland & Lewis, 2002). However, these developments have overlooked those with a greater degree of motor and cognitive impairment. Micro switches used to control sound and light displays have been documented as providing a means by which people with profound intellectual disabilities can increase their level of purposeful activity and exert some control over the stimuli that the environment can provide (Lancioni et al, 2001a). Learning to use the micro switch could also act as a focus for social interactions between the person with intellectual disabilities and staff. Previous studies using micro switches (for a review see Lancioni et al, 2001b) have linked the micro switch to a series of single auditory or visual stimuli. Developing computer based software that will provide a variety of responses and displays when controlled by a micro switch could provide just the activity that people with profound disabilities need to increase their level of activity, give them some control over environmental stimuli and provide a focus for social interaction with carers.

This study set out to

1. explore whether people with profound disabilities could learn to use a simple game controlled by a single micro switch and displayed on a normal computer monitor.
2. as a first step in examining social routines stimulated by playing the game, document what activities on the part of a tutor best facilitated the performance of the learner.

2. METHODS

2.1 Design of study

It was proposed to us a within subjects design to document changes over time in participants' behaviour and the help they received over repeated sessions with a switch controlled game displayed on a normal computer monitor. Evidence that participants could play the game would be an increase in the frequency with which they press the switch at the correct time as recorded by the score achieved on the game. Tutor activities would be considered to facilitate the learner's performance if they preceded correct switch pressing more than would be expected by chance.

2.2 Participants

Four men and three women aged between 24 and 46 years took part. They had profound disabilities but were described by their carers as having an appreciation of cause and effect. They also had sufficient vision, hearing and physical ability to press a micro switch and watch the computer monitor.

2.3 Software

The computer game displayed on a normal computer monitor was "Reign of the Flowerpots" (see Figure 1.) by RJ Games Ltd a San Francisco based company specialising in accessibility options for people with a wide range of disabilities http://www.rjcooper.com. Closure of the switch simultaneously caused a jet of water to emerge from a horizontally moving hose pipe and reversed its direction. If the jet of water hit a descending flowerpot the flowers would open in a spectacular display. If an un-watered flowerpot collided with the hose this triggered an entertaining visual and auditory signal. The speed at which the flowerpots descended could be increased as users became more familiar with the game.

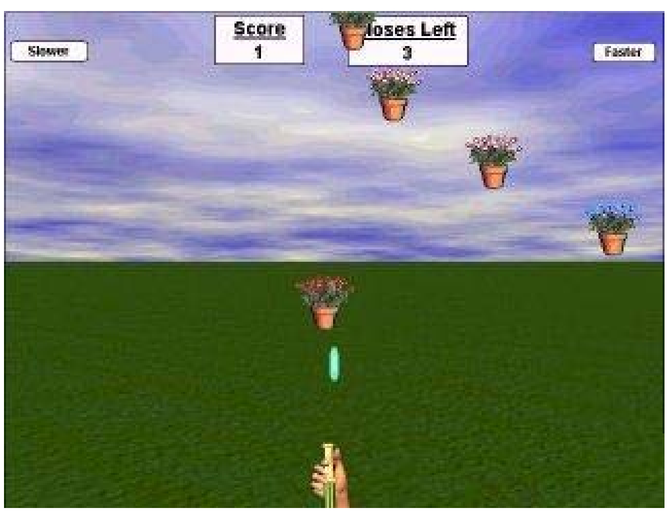


Figure 1. *Screen shot from Reign of the Flowerpots*

2.4 Data collection

Participants completed up to eight twice weekly sessions when they were given the opportunity to play the game by operating a large jelly bean switch positioned to suit their physical capability. A tutor (KA) sat next to them throughout the sessions and each session was recorded on videotape. The camera was positioned to view both the participant and the tutor. Tapes were analysed using OBSWIN for the help given by the tutor, use of the switch and duration of the participants' attention.

3. RESULTS

3.1 Could the participants learn to play the game?

For the learner it was essential that they attended to the screen, kept their hand on the switch and pressed it. Judging when to press the switch to score was difficult especially as some of the group took time to execute a movement sufficient to close the switch. The game turned out to be far too difficult for the learners so the scores recorded by the computer could not be reliably used as a measure of achievement. Alternative ways of evaluating the game had to be utilised. These are described below.

3.1.1 Participants' willingness to spend time in the test situation. Four people completed at least seven sessions and were happy to continue for the proposed session length of ten minutes. On the other hand, one indicated that she wished to finish after four minutes on her first session and withdrew from the study after completing only four sessions.

3.1.2 Time spent looking at computer monitor. All participants increased the percentage of time during the session in which they looked at the computer monitor (see Figures 1 and 2).

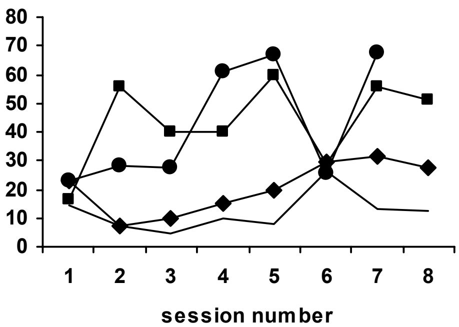


Figure 1. *Percentage of session spent looking at monitor for four participants*.

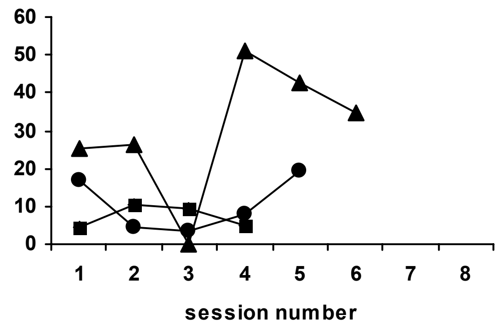


Figure 2. *Percentage of session spent looking at monitor for three participants.*

3.1.3 Rate of switch pressing. Of the six participants who completed five or more sessions, two had high rates of key pressing (20.33 and 15.96 per minute respectively) on their initial session (see Figures 4) whereas four had very low rates (1.18 – 2.6) and showed improvement over repeated sessions (see Figure 5).

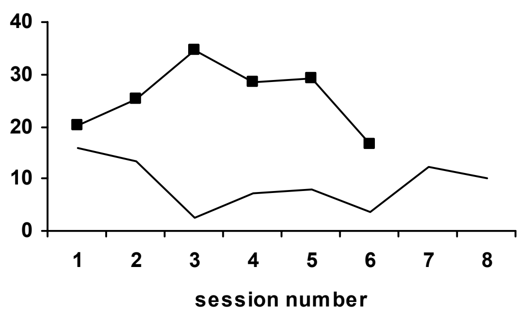


Figure 4. *Rate of key pressing each session for two participants*

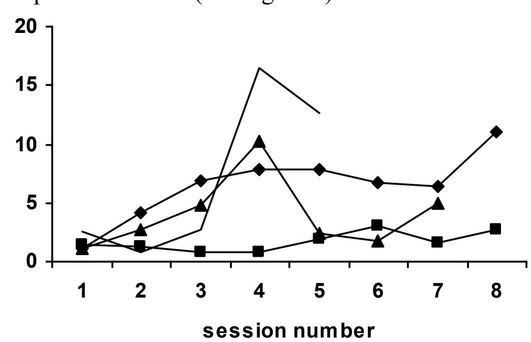


Figure 5. *Rate of key pressing each session for four participants*

3.2 What tutor activities promoted switch pressing?

Videotape analysis revealed the following actions taken by the tutor to promote staying at the computer, holding and pressing the switch and attending to the screen:

- *Verbal encouragement* (encourage)
- *Pointing at the screen* (point)
- *Pressing the switch herself* (press)
- *Placing the learner's hand on the switch* (put hand on)
- *Physically aiding them to press the switch* (press with T)
- *Telling the learner to stop injuring himself or herself or damaging the computer*

The tutor attempted to vary her strategies to suit the learner (egg trying to draw the learner's attention to the screen for those who were happy holding and pressing the switch; physically assisting them to press the switch if they failed to press it spontaneously). This is supported by an analysis of tutor activities that preceded switch pressing by the learner more frequently than would be expected by chance.

For all seven participants, there were at least two sessions when their switch pressing followed something the tutor did more often than would be expected by chance. For five of them, switch pressing followed the tutor's switch pressing significantly more often than would be expected by chance. Tables 1 and 2 illustrate this for two participants (C and M) who initially had low rates of switch pressing. Markov chain analysis was carried out on each session for each participant. Values of $\chi 2$ are given to indicate whether any of the five tutor strategies listed in the first row preceded the participant's switch pressing more frequently than would be expected by chance. Significant results are indicated in bold type: * $p<0.05$; ** $p<0.02$; *** $p<0.01$; **** $p<0.001$. If values of $\chi 2$ increase in later sessions this would indicate that switch pressing was becoming less random and more entrained to the behaviour of the tutor.

For two participants, verbal encouragement was also effective. One (Ra, see Table 3) was initially a high responder but the other was not. For another (R) who had an initially low rate of responding, his switch pressing followed switch pressing where he had been assisted by the tutor. (see Table 4). For none of the participants did this analysis suggest that their switch pressing was becoming less random and more entrained to the tutor with increasing exposure.

Table 1. *Tutor behaviours that preceded switch pressing by C.*

	Tutor action				
session	encourage	press	point	put hand on	Press with T
1		**4.14***			
2		**6.97*****	1.12	0.28	
3	0.12	0.11	2.37		
4	1.20	**6.29****	0.62		
5	0.01	**6.44****	0.91		
6	0.32	1.96	0.51		
7	2.73	2.39	0.78		
8	1.72	**11.72******			

Table 2 *Tutor behaviours that preceded switch pressing by M.*

	Tutor action				
session	encourage	press	point	put hand on	Press with T
1	0.92	**6.12****	0.64	2.68	
2	0.13		0.21		
3	1.96	0.09		2.73	
4	1.94	**10.24*****	1.29	0.80	
5	0.02	**13.60******			

Table 3. *Tutor behaviours that preceded switch pressing by Ra.*

	Tutor action				
session	encourage	press	point	put hand on	Press with T
1	0.75	**12.55******	1.24	0.89	0.83
2	**5.11***	**4.13***	0.21		
3	0.25	2.7		**5.76****	
4	0.96	**10.54*****			
5	0.35	4.18	2.25		
6	0.80	1.84			
7	**7.51*****	**17.07******	1.00		
8	**10.27*****	**5.64****	0.33		

Table 4. *Tutor behaviours that preceded switch pressing by R.*

	Tutor action				
session	encourage	press	point	put hand on	Press with T
1	0.26			**4.24***	**5.22***
2	0.81	1.33	1.10	0.17	**9.3*****
3	2.15				**16.22******
4	0.95			0.10	1.01
5	2.88	0.25			1.97
6	0.10	1.17	0.04	**11.71******	**4.89***
7	**5.30***	0.12			

4. DISCUSSION

The first aim of the study was to discover whether a group of people with profound disabilities could play a switch controlled computer game. Although key-workers of potential participants had seen the game and were asked to nominate people to take part in the study who had an appreciation of cause and effect, it soon became obvious that the game was going to be too difficult for all of those who took part. Rather than stop the sessions and disappoint both participants and day centre staff and in the absence of an easier game, it was decided to break down game playing into component skills and use shaping and reinforcement to increase the frequency of these component skills. The sub components of playing this particular game were attending to the screen, keeping their hand on the switch and pressing it at the right time.

All participants demonstrated an increase in the percentage of the session time in which they looked at the monitor. If not looking at the monitor they were looking at the switch or at the tutor, not surprising given that the tutor was a source of information about the situation as well as offering much valued social stimulation. This suggests that visual cues from the game either to indicate when to press the switch or to provide feedback, should be sufficiently salient to catch their attention.

Four of the participants also increased their rate of switch pressing. For two, their initial rate was high so the challenge for them was to press the switch at the right time. As the game was too difficult for them to discern cues signalling when to press the switch, it was essential that the tutor gave them this information. For all seven participants there were at least two sessions when their switch pressing did become a consequence of the tutor's activity be it verbal encouragement, pressing the switch herself or physical guidance. This could be seen as the rudiments of social routines or formats that Bruner (1975) originally described as developing between carers and infants. These were seen as essential steps in the development of the child's ability to communicate. The mutual ritual provided by various interactional games "scaffolds" the child's attempt, minimising the risk of failure. The lack of a trend towards more association between the switch pressing of the participant and the behaviour of the tutor with increasing exposure may be due to the design of the study. Participants spent a maximum of eight ten-minute sessions playing the game. Most studies of micro switch use involve considerably more exposure. For example, the two boys in the study by Lancioni et al (2001) used micro switch stimulation for ten minutes, four to six times a day for about twelve days.

As the cues provided by the game were inadequate to help the participants learn when to press the switch, their learning would have been facilitated by including a much simpler level to start with where each switch press controlled only one action, not two simultaneously and where more cues were given on the correct timing of the switch press.

The nature of this type of intervention inevitably means that only a small number of participants will be involved. This study included seven people who varied in age, gender and interests although they were very similar in ability. However, presenting group results is not always useful as opposing trends associated with different individuals can cancel each other out.

6. CONCLUSIONS

The study set out to see whether people with profound disabilities could play a game controlled by a single switch. The game was too difficult but all participants increased the time they looked at the monitor. For all participants, there were at least two sessions when their switch pressing followed something the tutor did more often than would be expected by chance. This could be seen as the rudiments of social routines or turn taking. For none of the participants did this analysis suggest that their switch pressing was becoming less random and more entrained to the tutor with increasing exposure. Participants varied in the type of action taken by the tutor that preceded their switch pressing.

When designing switch controlled games for this user group, visual cues from the game either to cue switch pressing at the right time or to provide feedback, should be sufficiently salient to catch the attention of the player. Each switch press should control only one action, not two simultaneously and more cues should be given on the correct timing of the switch press.

7. REFERENCES

D J Brown, N Shopland and J Lewis (2002), Flexible and virtual travel training environments, *Proceedings of the Fourth International Conference on Disability, Virtual Reality and Associated Technologies*, In Sharkey, Sik Lányi & Standen (eds), 181-188, Veszprém, Hungary, 18-20 Sept.2002.

J Bruner (1975), From communication to language, *Cognition*, **3**, pp. 255–287.

J A Clegg, P J Standen and J J Cromby (1991), Interactions between adults with profound intellectual disability and staff, *Australia and New Zealand Journal of Developmental Disabilities*, **17**, *4*, pp. 377–389.

G E Lancioni, M F O'Reilly, D Olivar and M M Coppa (2001a), Using multiple micro switches to promote different responses in children with multiple disabilities, *Research in Developmental Disabilities*, **23**, pp. 149–160.

G E Lancioni, M F O'Reilly and G Basili (2001b), Use of micro switches and speech output systems with people with severe/profound intellectual or multiple disabilities: a literature review, *Research in Developmental Disabilities*, **22**, pp. 21–40.

G E Lancioni, M F O'Reilly, N N Singh, D Olivar and J Groeneweg (2002), Impact of stimulation versus micro switch-based programmes on indices of happiness of people with profound multiple disabilities, *Research in Developmental Disabilities*, **23**, pp. 149–160.

S Landesman (1987), The changing structure and function of institutions: a search for optimal care environments, In *Living environments and mental retardation* (S Landesman & P Vietze, Eds) American Association on Mental Retardation, Washington.

P J Standen, D J Brown and J J Cromby (2001), The effective use of virtual environments in the education and rehabilitation of students with intellectual disabilities, *British Journal of Educational Technology*, **32**, *3*, pp. 289–299.

Towards eye based virtual environment interaction for users with high-level motor disabilities

R Bates[1] and H O Istance[2]

School of Computing, De Montfort University,
The Gateway, Leicester, UK

[1]*rbates@dmu.ac.uk*, [2]*hoi@dmu.ac.uk*

[1]*www.cse.dmu.ac.uk/~rbates*, [2]*www.cse.dmu.ac.uk/~hoi*

ABSTRACT

An experiment is reported that extends earlier work on the enhancement of eye pointing in 2D environments, through the addition of a zoom facility, to its use in virtual 3D environments using a similar enhancement. A comparison between hand pointing and eye pointing without any enhancement shows a performance advantage for hand based pointing. However, the addition of a 'fly' or 'zoom' enhancement increases both eye and hand based performance, and reduces greatly the difference between these devices. Initial attempts at 'intelligent' fly mechanisms and further enhancements are evaluated.

1. INTRODUCTION

Virtual Environments have much to offer motor impaired users such as entertainment, rehabilitation training, collaborative activities with users in remote places and the opportunity to experience a sense of place afforded by remote locations. A key aspect in realising this is the provision of interaction devices and techniques that are efficient in utilising the residual capabilities of the motor impaired users, whilst not requiring undue effort or imposing undue workload.

Eye gaze-based pointing devices have been used for interaction with 2D graphical environments for some time and can offer access to these environments for people who may have few other choices of interaction device. However, anecdotal evidence suggests that these pointing devices are an inefficient and unsatisfying means of interaction and manipulation due to the inaccuracy of eye-tracking systems, making direct interaction with environments difficult for disabled users who use these devices. Our previous work has suggested that much of this inefficiency and low user satisfaction with eye-based interaction can be greatly resolved by the use of supporting 'soft' devices to aid interaction (Bates 2000, Bates and Istance 2002).

The challenge now lies in utilising the knowledge and experience gained from 2D environments in enabling effective eye-based interaction with 3D virtual environments. An essential element in doing this lies in the provision of, and evolution to 3D, of the 2D 'soft' device enhancements shown to enable effective eye-based interaction in 2D environments.

2. PREVIOUS WORK

2.1 Eye-Gaze Interaction in 2D and 3D Environments

One of the goals of 3D virtual environments has been to enable users to apply the natural skills that they use in the real-world to their interaction with the virtual world. One of the advantages claimed for interaction techniques based on eye-gaze for 2D interaction is that these are more natural than the usual hand-based techniques, which involve a mouse or a joystick (MacKenzie et al. 2001, Jacob 1995, Jacob 1991). Also eye-based interaction in 2D contexts has been expected to be more efficient and faster as a means of pointing than other devices (Edwards 1998, Jacob 1995, Salvucci and Anderson 2000, MacKenzie 1992). Eye gaze is more direct as a user will first look at a target and then manoeuvre a pointer to it by means of a series of coordinated hand movements. Eye gaze should therefore be highly suited to interacting with 3D virtual

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

environments with its promise of enhanced naturalness and efficiency, if the limitations of eye-based interaction found in the context of 2D interaction can be overcome.

An experiment where hand-based pointing, using a joystick, and eye-based pointing were compared for targets of different sizes within a virtual environment has been conducted (Asai et al 2000). Here they reported that eye-based pointing was 10 times faster on average across all target sizes than hand-based pointing. They note though that lack of familiarity with the joystick may have been responsible for some of the longer times recorded in the hand-based condition.

A significant example of the early use of eye-gaze in interacting with 3D virtual worlds was the 'self-disclosing' display (Starker and Bolt 1990). Eye-gaze was used to assign an 'index of interest' to objects in a virtual world presented on a desktop display. The system responded dynamically to changes to the assigned indices of interest by adapting a spoken commentary about the world to reflect either a particular interest in one object, or an interest spread across a group of objects.

This technique was subsequently used as a gaze-based interaction technique within a virtual environment (Tanriverdi and Jacob 2000). This entailed that objects in a virtual environment receiving a sufficiently high index of interest increased in size and revealed their internal structure. This is not a 'zoom-in' of the observer towards the object of interest, but rather a 'zoom-out' of the object towards the observer. Tanriverdi and Jacob compared the performance of this technique with that of a hand-based pointing technique for a task which involved selecting objects in an environment to discover which contained various target stimuli. These targets could either be reached in the hand pointing condition while the user was stationary, or by the user moving 5 to 15 inches forward. A performance advantage in favour of gaze-based pointing over hand-based pointing was found, but only for targets which required the subject to move forward. The gaze pointing condition enabled selection of the same targets without the user having to move. This presumably afforded some advantage to the gaze condition regardless of any inherent difference between eye and hand as a pointing modality.

A similar experiment was conducted (Cournia et al 2003) but used the 'ray-casting' interaction technique for both hand and eye conditions. There was no 'zoom-out' for any target, only a fade to reveal inner structure. They found a significant difference in favour of *hand-based* pointing for targets at distances where Tanriverdi and Jacob had found advantages for eye-based pointing. These two pieces of work suggest that the benefits of eye-based pointing in virtual environments are contingent on the apparent size of target (assuming targets located further away are smaller) and on the interaction technique used.

2.2 Enhancing Eye-based Pointing in 2D and 3D Environments

Interaction in 3D environments can broadly be characterised as object manipulation, navigation and application control (Hand 1997). Zooming or flying in towards an object can be seen both as a navigation technique and as an object manipulation technique. Temporarily zooming-in on an object of interest to select it makes it easier to select objects with an inaccurate pointing device (Bates and Istance 2002), although it is important to be able to zoom back out to the original position to prevent loss of context and orientation. Without the return to the original position from where the zoom action was initiated, zoom becomes a fly navigation technique ('fly where I point' or 'fly where I look').

'Intelligent flying' can utilise a similar technique to the 'index of interest' where initiating a 'fly' action assumes the target to be that object with the highest index of interest. Additionally the fly can stop at a reasonable distance in front of the assumed object of interest so that it does not fill too much of the visual field, or indeed, to ensure that the user does not fly straight through the object. Moving the gaze point during the fly can either effect small corrections to the flight path or indicate the intended object to fly to is not that which has been assumed by the system.

3. DETERMINING THE BENEFITS OF FLYING TO 2D OBJECTS

3.1 An Experiment in 2D

An experiment was conducted in a complex 2D GUI test environment to measure the effect of providing a basic 'fly', or zoom, enhancement (Bates and Istance 2002). Hand and eye based pointing devices, or hand and eye mice, with and without the fly enhancement were used to manipulate GUI objects of four angular sizes based of the angle the objects subtended from the eye of the user. These ranged from 0.3^0 to 1.2^0. Interaction typically lasted for 20 minutes and incorporated 150 test tasks for each of the 6 users in the experiment. The objective efficiency (based on time and quality of interaction metrics) and subjective user satisfaction (based on ratings of workload, comfort and ease of use) of the manipulation were measured. These objective metrics were derived from the ISO9241 part 11 standard and the ESPIRIT MUSiC metrics method and previously validated. The subjective metrics were obtained from a set of rating scales, similar to the NASA TLX. For this experiment the basic fly enhancement was under the full control of each user via micro switches rather than under intelligent software control in order to determine at what distance of fly into the environment the users opted to stop the fly and start manipulation.

3.2 2D Flying Experimental Results

Fig. 1 shows the task efficiency metrics for the hand mouse, the eye mouse without fly and the eye mouse with the fly enhancement for each of the target size categories. The results for the eye mouse without the fly-enhancement (Fig. 1, 'Normal' mode) showed near-unusable efficiencies for the smallest object at 18% task efficiency. Efficiency increased, as expected, with increasing target size. It was notable that even with the largest object size that the eye mouse efficiency was less than the hand mouse efficiency at the smallest target size.

Examination of the results of the eye mouse enhanced with the 'fly' soft device (Fig. 1, 'Fly' mode) showed dramatic increases in efficiency for all object sizes, with the non-fly performance on the largest target being equalled by the 'fly' efficiency on even the smallest target size. Note that this improvement includes the detrimental overhead of manipulating the 'flying' during interaction; with the measurement of efficiency including penalties for time spent controlling the fly and errors in controlling the fly distance.

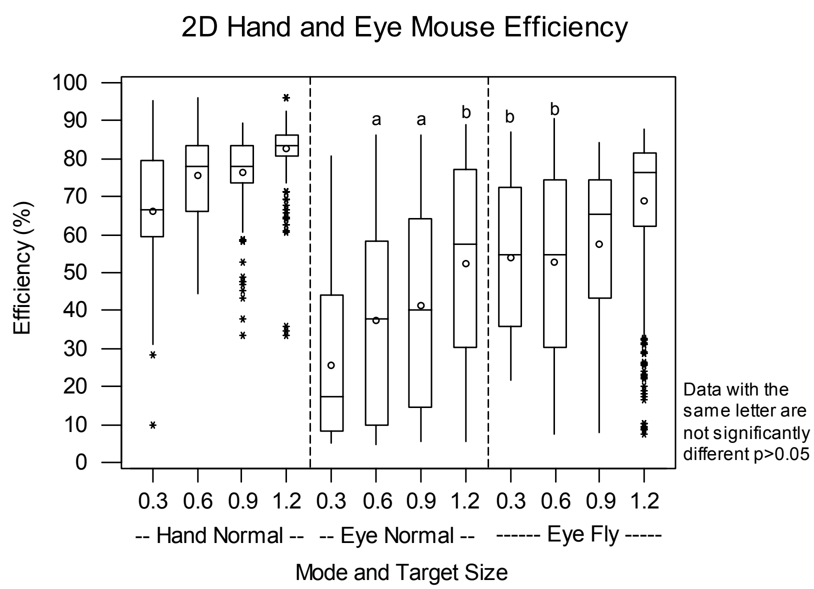


Figure 1. *Hand and eye mouse efficiency and the effect of object size and flight in 2D.*

Examination of the efficiency results by object size showed a clear relationship between object subtended angle and use of fly, with fly being increasingly used as object angles decreased (Table 1). For eye-based pointing the effective object size after flying was near constant at 1.7 degrees for all objects. These results gave the basis for calculating a 'smart fly' automatic stopping distance in front of an object for the device.

Table 1. *Eye mouse effective flown object subtended angles at 2D fly stop point.*

Original angle subtended	Enhanced eye 'flown' angle
0.3°	1.6°
0.6°	1.6°
0.9°	1.8°
1.2°	1.7°

Examination of the satisfaction ratings (Fig. 2) showed high levels of workload for the eye mouse, although with a slight decrease in rating for the enhanced mode. This was encouraging as the decrease was in spite of the overhead of additional workload controlling the fly. As expected, comfort levels remained unchanged as there was little that would affect comfort during use. There was a slight increase in device ease of use for the fly condition. It is likely that greater efficiencies and larger decreases in workload and increases in comfort and ease of use will arise from an automated 'smart fly' facility where the user does not need to actively control the fly.

Figure 2. *Hand and eye mouse satisfaction and the effect of flight in 2D.*

In conclusion, the provision of a zoom or fly enhancement greatly increases the efficiency of eye-based pointing on a 2D GUI, without adversely affecting subjective ratings of comfort or workload. Furthermore, when the user has full control over the extent of the zoom, it is used such that targets of originally different sizes all subtend approximately 1.7^0.

4. DETERMINING THE BENEFITS OF FLYING TO 3D OBJECTS

4.1 An Experiment in 3D

An experiment was conducted in a virtual environment to examine the extent to which hand-based and eye-based pointing would benefit from a similar fly enhancement to that previously examined with the 2D zoom enhancement. The hand-based and the two eye-based conditions (one with fly enhancement and one without) all used ray casting as the interaction technique, as did Cournia et al. previously.

Unlike previously reported work, which has used an immersive head mounted display, this experiment was conducted in a reality centre located at De Montfort University, equipped with passive stereoscopic images across an 8-metre wide 150^0 cylindrical screen. In all conditions, users were seated 6 metres away from the curved screen. A desk mounted SMI RED eye tracker was used for eye-based pointing, and a desktop mouse was used for the hand based pointing in order to enable direct comparison with the 2D environment. Six users took part in the experiment and a within-subjects design was used.

4.2 Experimental Task

The task was to select one of a group of virtual students in a virtual lecture theatre. The required target to select was indicated by a hat appearing on the student. Students at the back of the lecture theatre subtended the smallest visual angle in the experiment. The students on the four rows subtended the same four visual angle sizes as in the first 2D experiment.

An illustration of the experiment in the VE is shown (Fig. 3). Here the upper left frame shows the VE before a fly is invoked. The end of a vector from the user's eye through the VE can be seen as a grey cube in the middle of the picture, with the desired target object signified as the person wearing a hat. The upper right frame shows the subject 'flying' rapidly (12.5m/s) toward the target object. Finally, the lower frame shows the target object being selected with the hat flying off and away from the target. Interaction in the experiment typically lasted for 20 minutes and incorporated 144 test tasks for each user. As before, the objective efficiency and subjective user satisfaction of the manipulation were measured.

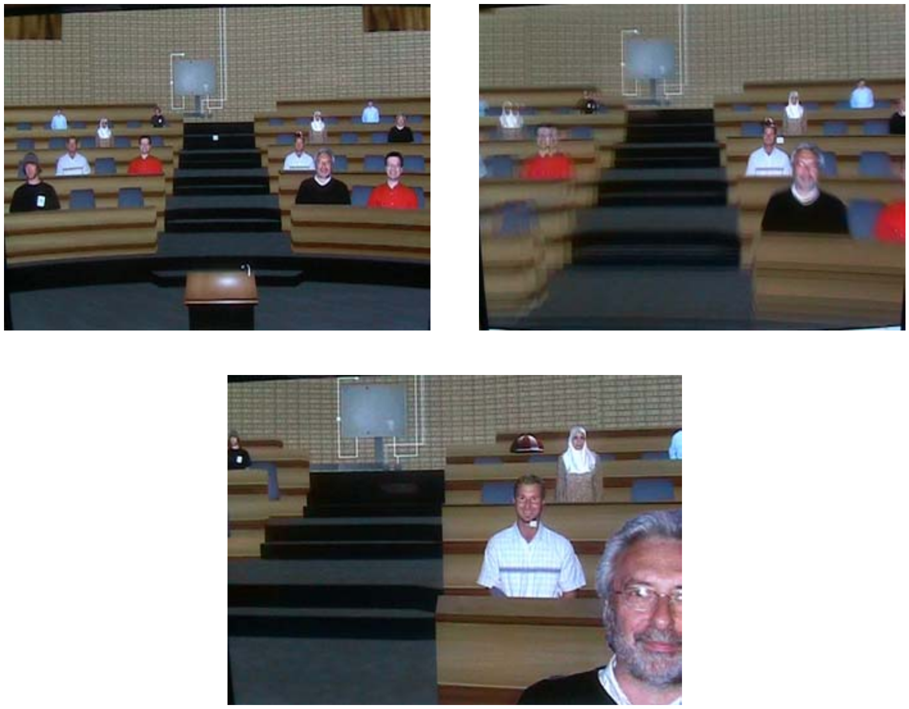


Figure 3. *Flight and object selection in the VE.*

4.3 D Flying Experimental Results

Following the procedure of the 2D experiment, the results were broken down by target size visual angle for all devices. For the hand device, objective task efficiency and subjective user satisfaction (Figs. 4 and 5, left hand and centre sets of data only) showed that device performance was highly dependent on target size, with poor performance for the smaller object sizes. Comparing the hand mouse efficiency in a 2D environment (Fig. 1) to that in the 3D environment (Fig. 4) showed a marked drop in performance for all object sizes in the 3D environment. Enhancing the hand mouse with the 'fly' soft device (Fig. 4, 'Fly' mode) showed large and significant increases in efficiency for the smallest three object sizes, with no significant improvements for the largest object size, where fly was rarely used.

The subjective workload ratings (left bars in each graph sub-section Fig. 5) showed increased workload and lower ease of use in the 3D environment compared to the 2D environment (left bars in each graph sub-section Fig. 2). It was notable that observed workload ratings were lower, and the ease of use ratings were higher in the 3D environment (centre bars in each graph sub-section Fig. 5) when the fly enhancement was used with hand-based pointing.

The efficiency results for the eye mouse without the fly-enhancement (Fig. 6, 'Normal' mode) showed extremely low efficiencies for the smallest objects, with efficiency increasing, as expected from the 2D results, with increasing target size. As with the 2D results, the eye mouse showed lower performance than the hand mouse, although the differences were considerably reduced between the devices in the 3D environment. There is no difference for the smallest target size, but as target size increases hand pointing outperforms eye pointing. This difference is similar to that found by Cournia et al quoted earlier. With the fly enhancement (Fig. 6, 'Fly' mode) the efficiency for all target sizes was increased, with object size now having a lesser effect on efficiency, and the eye mouse achieving near parity with the hand mouse. Examination of the satisfaction ratings (Fig. 7) showed improvement for all ratings for the enhanced mode. As with the hand mouse, it was notable that the fly enhancement reduced the observed workload ratings and increased the observed ease of use ratings in the 3D environment.

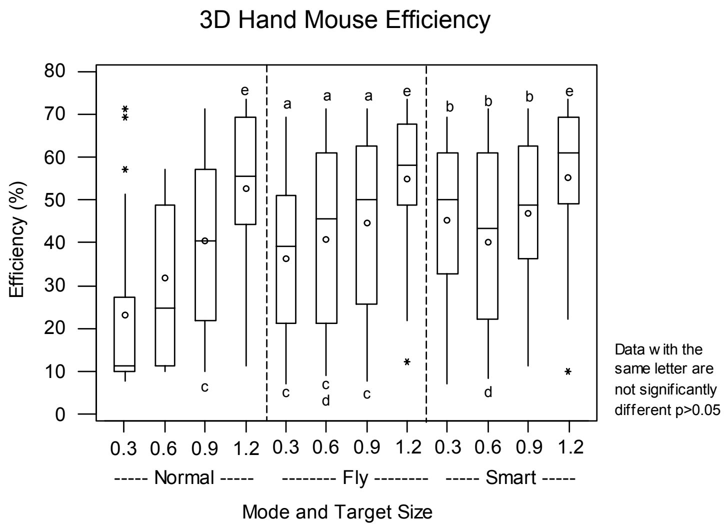


Figure 4. *Hand mouse efficiency and the effect of object size and enhancement in 3D.*

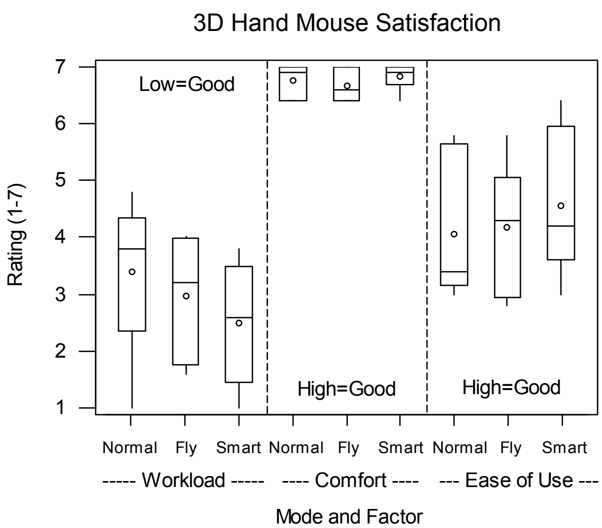


Figure 5. *Hand mouse satisfaction and the effect of enhancement in 3D.*

As before, the efficiency results gave the basis for calculating the stopping distance in front of an object for both of the devices (Table 2). Unlike the 2D condition, there was less consistency in the level of target magnification at the point when the target was selected.

Table 2. *Eye mouse effective flown object subtended angles at 3D fly stop point.*

'Normal' angle subtended	'Fly' hand effective angle	'Fly' eye effective angle	'Smart' hand effective angle	'Smart' eye effective angle
0.3°	1.3°	2.2°	1.3°	2.1°
0.6°	1.4°	2.6°	1.4°	2.2°
0.9°	1.5°	3.3°	1.5°	2.2°
1.2°	1.9°	3.0°	1.6°	1.9°

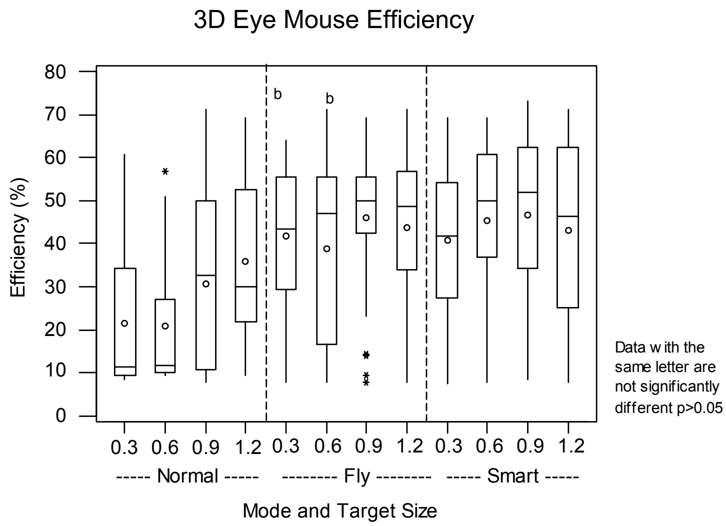


Figure 6. *Eye mouse efficiency and the effect of object size and enhancement in 3D.*

Figure 7. *Eye mouse satisfaction and the effect of enhancement in 3D.*

5. DETERMINING THE BENEFITS OF INTELLIGENT FLYING

5.1 An Experiment with 'Intelligent' Flying

Based on 3D experimental results, a 'smart' stopping distance was set at a distance where the visual angle of an object would subtend 2.4°; a compromise to give greatest ease of manipulation without overly enlarging objects and potentially disorienting the user, but also flying sufficiently close to give ease of selection. To do this the fly enhancement was modified to monitor the point of interest in the environment, and the required fly distance toward target objects based on their apparent size, with the fly automatically being stopped when the object subtended a 2.4° visual angle. After manipulation, the user returned to the original starting point by initiating an automatic fly back. This gave the first elements of an 'intelligent' fly, with the fly automatically stopping at an optimal distance, and then returning to the pre-flight position. In the same manner as before, a trial was conducted with the intelligent fly enhancement using hand and eye pointing.

5.2 'Intelligent' Flying Experimental Results

The results for the 'intelligent fly' hand and eye mice were appended to the 'normal' and 'fly' results to aid comparison between the three modes of operation (Figs. 4 to 7). Overall, device efficiency was essentially unchanged from the basic 'fly' mode, showing that although test subjects tended to fly closer to objects with the eye mouse (Table 2) than the estimated ideal distance, there were no performance benefits from doing so. In addition, the intelligent fly suggested reduced workload is possible in comparison to the basic fly mode.

6. CONCLUSIONS

The work reported has demonstrated that the benefits of enhancing eye pointing by zoom previously demonstrated in 2D interaction are also apparent in 3D interaction. Our results comparing eye and hand pointing in virtual environments accord with those reported elsewhere, and show a performance advantage for hand based pointing. Our work shows that this benefit is only apparent for larger target sizes, however, when a zoom or 'fly' enhancement is provided the performance levels of eye based pointing increase to a similar level to that of hand based pointing. We have not been able to demonstrate the same consistency of zoom that was apparent in the 2D environment. Our initial attempts to go further and introduce a degree of 'intelligence' have indicated some success. The limited but promising results suggest that more effort is required to add further intelligence to interaction to gain performance benefits.

The addition of an intelligent control based on optimal object subtended angles is currently under investigation and is expected to further enhance performance with eye based pointing in 3D environments. In this way, the naturalness and efficiency benefits offered by this modality should be realised to a greater extent. This work forms part of, and contributes to, the efforts of the Communication by Gaze based Interaction (COGAIN) fp6 Network of Excellence.

7. REFERENCES

K. Asai, N. Osawa, H. Takahashi, Y.Y. Sugimoto, S. Yamazaki, M. Samejima and T. Tanimae (2000) Eye Mark Pointer in Immersive Projection Display, *Proc. IEEE Virtual Reality 2000 Conference,* p125.

R. Bates (2000) Using Soft Devices as Enabling Translators - Standard GUI Interaction For All! *Proc. Human Computer Interaction,* **2**, pp53-54.

R. Bates, and H.O. Istance (2002) Zooming interfaces! Enhancing the performance of eye controlled pointing devices, *Proc. ASSETS 2002, The Fifth International ACM SIGCAPH Conference on Assistive Technologies*, July 8 - 10, 2002, Edinburgh, Scotland.

N. Cournia, J.D. Smith and A.T. Duchowski (2003) Gaze- vs. Hand-Based Pointing in Virtual Environment, *in Proc. SIGCHI 2003 (Short Talks & Interactive Posters)*, April 5-10, 2003, Ft. Lauderdale, FL.

G. Edwards (1998) New Software makes eye tracking viable: You can control computers with your eyes, in *Proc. CSUN Technology and Persons with Disabilities Conference*, California State University Northridge.

C. Hand (1997) A Survey of 3D Interaction Techniques, *Computer Graphics Forum,* **16**(5) pp269-281.

R.J.K. Jacob (1991) The use of eye movements in human-computer interaction techniques: What you look at is what you get, *ACM Transactions on Information Systems* **9**, *3,* pp152-169.

R.J.K. Jacob (1995) Eye tracking in advanced interface design, *in Advanced Interface Design and Virtual Environments*, pp258-288, Oxford University Press.

I.S. MacKenzie (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, **7**, pp91-139.

I.S. MacKenzie, T. Kauppinen and M. Silfverberg (2001) Accuracy Measures for Evaluating Computer Pointing Devices, *Proc. of CHI2001*, ACM Press.

D.D. Salvucci and J.R. Anderson (2000) Intelligent gaze-added interfaces, *Proc. of CHI 2000*, **2**, ACM Press, pp273-280.

I. Starker and R.A. Bolt (1990) A Gaze-Responsive Self-Disclosing Display, *Proc. CHI '90*, ACM, pp3-9.

V. Tanriverdi and R.J.K Jacob (2000) Interacting with eye movements in virtual environments, Conf. Human Factors in Computing Systems, *Proc. SIGCHI conference on Human Factors in computing systems*, The Hague, The Netherlands, pp265 – 272, ACM Press.

Design, development and manufacture of novel assistive and adaptive technology devices

S J Battersby[1], D J Brown[1], P J Standen[2], N Anderton[2] and M Harrison[3]

[1]Department of Computing and Mathematics, The Nottingham Trent University, Burton Street, Nottingham, ENGLAND

[2,]Division of Rehabilitation and Ageing, QMC, University of Nottingham, ENGLAND

[3]The Portland Partnership, Portland College, Mansfield Road, Nottingham, ENGLAND

steven.battersby@ntu.ac.uk, david.brown@ntu.ac.uk, P.Standen@notingham.ac.uk, mharrison@portland.org.uk

[1]*www.ntu.ac.uk,* [2]*www.nottingham.ac.uk,* [3]*www.portland.ac.uk*

ABSTRACT

The aim of this research is to design, develop, evaluate and manufacture an assistive/adaptive computer peripheral to facilitate interaction and navigation within Virtual Learning Environments and related learning content for people with physical learning and disabilities. The function of the device will be software specific; however the most common primary functions are those of selection, navigation and input.

1. INTRODUCTION

This research project has been developed within the Portland Partnership Project which has been funded under the first round of the European Equal Call. The holistic aims of the project are to develop Virtual Learning Environments (VLE) to teach basic skills to young people with severe physical and associated cognitive disabilities.

The Portland Partnership is led by Portland College (PC). The VLE is being developed by North Teeside University (Tee), with the embedded Flash based learning materials being developed by Cambridge Training and Development Ltd (CTAD). The aim of the research described in this paper is the design, evaluation and manufacture of a switch based device to provide control (scanning and selection) of the VLE and embedded learning materials by a group of people with a wide and heterogeneous range of physical and cognitive abilities.

This development of the assistive/adaptive device is being carried out in a partnership approach itself between the Nottingham Trent University (NTU), Nottingham University (NU) and Traxsys. These partners have the following roles in the design, development, evaluation and manufacture of the device:

NTU: Parallel Products Review of existing switch based devices to identify potential design solutions; User Group and Usability Group work to drive the contribution to design by people with a physical and cognitive disability and professionals in the field; Usability Study to inform the development of the Design Specification Document; Development of Automatic Data Capture Software for baseline and prototype evaluation studies; Prototyping of the assistive/adaptive device.

NU: Baseline Evaluation of existing switch based devices with prototype basic skills software developed by CTAD; Evaluation of Prototype with the same software for comparison with baseline; Identification of Modifications to prototype device based on comparative analysis, and Recommendations to Manufacturers.

Traxsys: Contribution to Usability Group and Design Specification Document via experience and manufacturability metrics; Manufacture of the assistive/adaptive device.

It is intended that the device will not only facilitate navigation (scanning) and interaction with the VLE and Flash based learning materials, but also contribute to the assessment of the user group by being mountable on other hardware and allow the determination and assessment of their individual switch based needs. The feasibility of the development of a bespoke piece of software to assess the switch based needs of

 ISBN 07 049 11 44 2

all learners is currently being investigated. The target population represents a heterogeneous group of users, with the definition of a universal design unlikely. Where the best design solution identified and manufactured as a result of this research does not fit a learner's abilities, the assessment software linked to a contemporary database of other existing solutions could be used to match a best fit.

It is worth noting that that the choice of Macromedia Flash to develop the embedded learning materials raises several issues. The first is that it eases the development of the data capture software to automatically capture evaluation results, significantly aiding the task of the evaluation researcher. The second is that its use as a development tool has shown positive results in motivating the learning activities of young people at risk of social exclusion and that its presentational format (animation driven) is ideal for young people with a cognitive disability who will have associated literacy difficulties (Brown et al, 2002).

The project is now at the stage where the Parallel Products Review is complete, the User Team and Usability Team are running and providing design assistance, advice and storyboards, The Baseline Assessment of existing devices with prototype CTAD software is proceeding and the Usability Study is now being collated from complete User, Task and Environmental analyses. In additional to these design components, the advice from device manufacturers Traxsys will complete the Design Specification.

2. DEVELOPMENT METHODOLOGY

The development methodology is based on one developed and evolved over a series of earlier related studies (Hall, 1993; Brown et al, 1997; Soos, 1998; Lannen et al, 2002 and Standen et al, 2002). Its application to this project is now described.

2.1 Parallel Products Review

Evaluation of existing technologies has been undertaken to provide a Parallel Products Review to determine the reusability of existing solutions in the research. The properties of an extensive range of switch based devices should be examined against the requirements of control (scanning and input), cost and the Portland Partnership Technical Specification. The results of this review will contribute towards the Design specification Document.

2.2 Usability Study

A user-sensitive design will be employed to meet the needs of users and beneficiaries. In user-centred design, product developments are driven from user requirements rather than from technological capabilities. Central to this design process is usability, a crucial factor in the production of successful human-computer interfaces. Usability is defined in ISO 9241, part 11 as:

> *"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."*

According to the ISO 13407 standard (human centred design processes for interactive systems), the key activities in user-centred design are:

> *"Understanding and specifying the context of use, specifying the user and organisational requirements, producing designs and prototypes and carrying out a user-based assessment."*

An iterative process is employed with the cycle of activities being repeated until the design criteria have been attained. The first activity can be achieved by conducting a Usability Context Analysis (UCA), which involves the following stages:

- User analysis: to describe physical, cognitive and perceptual abilities of the user group. Physical Ability: Quick Neurological Screening Test II.
- Task analysis: to identify the major productive goals, which a user can achieve, using the product.
- Environment analysis: to describe the organisational, technical and physical factors of the environment in which the product will be used.

2.3 Baseline Evaluation: Within Subjects Design

A within subjects design is being utilised to collect performance data on existing and devices, which can be compared in later analysis with new prototype devices to identify performance enhancement. To reduce the effect of learning on performance, participants will be allowed several practice sessions with the currently used devices before data collection starts. Data collected at each stage will be used to inform the design of the final prototype or any subsequent alterations.

Data will be captured automatically using the bespoke software capture suite designed and currently implemented at NTU. This will considerably speed up, and make more accurate, the process of data capture. This in turn will allow a greater amount of data to be captured for comparative analysis.

2.4 Data Collection method

Data collection will take place within Portland College. The Evaluation Researcher will sit alongside participants to give assistance and encouragement. The order in which they will work through the VLE will be the same for each participant starting with those environments which require fewer functions from the control devices.

Each session will be recorded on videotape, the camera positioned to view both the participant and the researcher sitting next to them. The videotapes will be analysed using a method established in an earlier author led study which yielded measures of help given by the researcher, user engagement, unproductive actions and user comments (Standen et al, 2002).

The researcher will also keep a diary to record any other information that might be useful but that would not be picked up by video analysis or the bespoke data capture software.

3. PROJECT PROGRESS

3.1 Project Management

A multi-disciplinary team was formed to advise and monitor the project development. To aid this process a website (Portland.isrg.org.uk) has been developed to provide an on-line reference project management and dissemination tool. The team includes a design engineer, human-computer interaction and special needs experts, a usability specialist, learners, an occupational therapist and a representative from industry (Traxsys).

3.2 Parallel Products Review

A survey of existing input and assistive/adaptive devices has been carried out to identify devices, features and components that will inform design and reusable components for the project. This, in conjunction with an over-arching Technical Specification provided by Portland Partnership, has raised the possibility of the development of an additional software solution to assess all learners' switch-based needs and to point to a variety of existing best-match solutions, where the device developed by this research does not fit their needs. This eventuality is possible due to the heterogeneous nature of the target population. The conclusions from this review were as follows:

Essentially all assistive switches are the same, really only differing in ergonomics and aesthetics. In order to interface with a computer system successfully, the switch needs to be used in conjunction with a suitable connector.

The Technical Specification provided by Portland Partnership has dictated that this connector be of the USB type. This factor, considered in conjunction with the requirement of being compatible with existing VLE and learning software developed in the project thus far, suggests the use of either a Don Johnston pro 5, Sensory Software JoyCable or Joybox Interface.

Whilst the Don Johnston pro 5 can be used without the instillation of an additional software solution, in comparison with the Sensory Software Joybox and JoyCable, it is not as adaptable. Recognised by windows as a switch controlled USB joystick, the Joybox and JoyCable devices can be used with any software that supports joystick input extending functionality.

The Joybox allows up to 12 switches to be connected via the USB port (see Figure 1), giving an advantage over the JoyCable. However, the JoyCable could possibly provide the optimal solution, dependant upon the outcome of the device's input requirements; i.e. 2 switches or less.

For software that does not support joystick input, a driver is provided to allow the switches to control the keyboard or mouse. This allows the switches to be used with any Windows software (and Windows itself), thus resulting in full switch interface functionality.

As previously stated all switches are very similar and in function provide an output in the form of a true or false signal dependent on the switch's state. At base level the technology used in all switches is the same. On triggering, an electric current is shifted to another circuit, providing a signal state.

The main differences are factors such as size, shape, colour and physical feel. Some developers have also added additional functionality such as the ability to alter trigger activation parameters. Another popular trend is the provision of auditory feedback.

At this stage it has been decided that a sample of approximately 10 switches should be tested with the users in order to obtain a good baseline evaluation of present switch technology. This outcome forms deliverable 1 of the project to inform the Product Design Specification Document in terms of reusability of existing design solutions and features.

Figure 1: *Sensory Software JoyBox*

3.3 User-based Development Team

A user-based development team has been formed to work in conjunction with the design engineer. The team consists of around six sixteen – nineteen year olds spanning four main levels of ability (pre-entry to prevocational), with additional potential access to users within the Portland College's Employment Department. Those within the Employment Department have less disability and more life experience and thus provide good additional user representation, however the main focus is to address the greatest challenge - "include the excluded".

These users will help to inform the deign of the adaptive/assistive device through consultation using an interactive storyboarding method first developed by the team for use in the design of VLE-based life skills learning software (Brown et al, 1999) and enhanced in recent studies and this research to develop navigation devices for the use of VLE by people with a cognitive disability (Standen et al, 2002). This methodology places users at the centre of a rapid prototyping development cycle that allows them to effect changes to the device design and quickly see the results of such action. This rapid turn around is important for people with a cognitive disability in order to maintain their interest and active involvement in the design process. Figure 2 shows this process in action for the development of the navigation device for people with a cognitive disability using VLE in a recent EPSRC sponsored study.

Figure 2: *Development of interactive storyboards for EPSRC Study*

The interactive storyboarding process for the switch-based device developed within this project started with an explanation of the storyboarding process using 3D based CAD. This explanation was given to the User

Team by the Design Engineer (NTU) and Portland Partnership Project Manager (PC) demonstrating 3D imagery in combination with manipulation of the real life counterpart object.

A mug (large cup) was held up in front of the User Team and rotated whilst its counterpart 3D computer generated image was rotated simultaneously on screen. Members of the User Team were then asked to choose their favourite colour. These colour settings were then applied to the electronically generated mug to illustrate the relationship between the two sets of objects in a concrete way. The User Team could then readily understand how the interactive storyboards represented real world objects and how they could contribute to modifying their properties to express their own design preferences.

The User Team was then shown a range of commonly used switch-based devices to gain their feedback on design preferences (on preferred colour, shape and textures). Figure 3 illustrates the method of describing the iterative prototyping process to the User Team, which contributes to Deliverable 2 to inform the Product Design Specification Document.

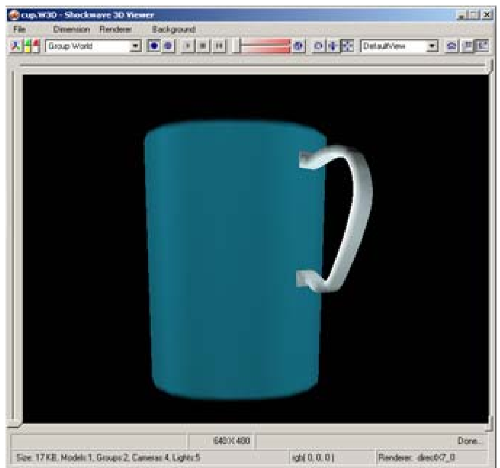


Figure 3: *User Team Design Preferences*

3.4 Usability Team

The Usability Team has a dual function within the project. Its first main task is to provide a steering committee (meeting six times) to effectively manage the project, monitor deliverables, provide dissemination and provide expert advice to contribute towards the Design Specification Document. To this end the overall Portland Partnership Manager (PC) provided visioning storyboards of the switch based assistive/adaptive device based on an expert knowledge of the target user group abilities, the Portland Partnership Technical Specification and a situated analyses of the environmental factors and tasks expected to be performed. These storyboards contribute to deliverable 3, as a component of the Design Specification Document.

3.5 Baseline Evaluation

Twenty six (26) students have been recruited to the baseline evaluation study who experience a range of physical and cognitive abilities described by the User Analysis (section 3.6.1). This evaluation study is currently being carried out by the Evaluation Research (NU) using the following commonly used devices (and also shown to the User Team to examine their design preferences): Big Red, Standard Jelly Bean, Spec, Palpad, The Handy Switch (Wobble Switch) and Churchill.

The baseline evaluation study (Deliverable 4) will contribute towards the Design Specification Document in two ways; by providing data for comparative analysis with the final prototype developed, and by the identification of usability problems users experience with currently existing and commonly used switch based devices when using the Flash based learning software developed by CTAD and embedded within the VLE. This evaluation study is using the bespoke designed data capture software and an established set of usability metrics.

3.6 Usability Study

3.6.1 User Analysis. Thirteen (13) male and thirteen (13) female participants, aged 16-22 years at recruitment, consented to take part in all phases of switch evaluation. They completed verbal Intelligence Tests (BPVS) and Motor Skills Assessment in QNST2. Most participants have limited vocabulary and varied methods of communication including speech, eye gaze, head movements and the use of communication devices. Most participants would be able to use a switch with one or both hands but are limited in reach, strength and motor control. Therefore position, size and texture of the switch developed needs to be explored.

3.6.2 Task Analysis. Given that a tutor has set up the session, the tasks that a student will have to perform within the VLE and embedded learning materials, with a relevant degree of support from the tutor, are:

- Log in to the VLE using relevant input device.
- At the earliest project development milestones (M1-3), when they click on 'learn' they will go straight into the activity (no choice).
- At later milestones (M4-6), they will have a choice of 2 items (if set up by the tutor in this way, although the tutor may choose a linear presentation of the materials) e.g., a video (to provide context) and an activity (learning object, which will include a test game at the end, with the score being passed back to the VLE database). The video and the activity will be represented on screen by thumbnail images or icons.
- At later milestones (M7-8), they will have a choice of 4 items.
- At Entry level, the VLE may appear more like a traditional VLE with a wider choice of items.

In terms of how these choices are presented, there is a menu presented to the students. For the early prototypes CTAD have developed a "cheese" icon as a means of selection as an interim measure. As their understanding deepens of the tasks to be performed within the embedded learning materials and how these are linked to the VLE, the use of this icon will need to be reviewed. It may be that the cheese icon is not the best metaphor as it is perhaps not versatile or flexible enough. It may be better to have a thumbnail or other icon for each activity; however, there is a need to balance this against the fact that learners have liked using the cheese.

At the end of a session using the VLE, it is important for the learner to have a summarized printout as a hard copy record of what they have achieved in the session (as well as a progress indicator in the VLE itself).

In conclusion, the task analysis has revealed that students will need to use a combination of scanning and selection (with choice) to navigate and interact with the VLE and embedded learning materials.

3.6.3 Environment Analysis. There are a range of technical and physical factors that form the environmental analysis. Technical factors that will exert an influence on the Design Specification Document have been supplied in the form of the Portland Technical Specification Document and can be summarized as following to be:

- wireless
- compatible with existing hardware (Discover switch, Joy Box, standard USB PC *etc*.) or better than this hardware in terms of functionality
- compatible with existing software and also any materials produced by the Portland Partnership
- potentially be used as an assessment device (in terms of being adjustable in specification (surfaces, mounting *etc*.)
- flexible in terms of physical attachments
- accompanied by a system which allows for the base unit to be moved within a specified range (45 cm)
- capable of adhering to the technical specifications as advised by the technical group of the Partnership.

Physical factors identified within the environment that may exert an influence on the Design Specification Document are:

- Many users use wheelchairs, and hence their switches will be used on wheelchair trays, rather on a desk. This may exert constraints of size and fixture abilities on the device.
- The time of day at which the switch and VLE will be used, as many students are more tired at end of day and the device may have to attempt to compensate for this.
- The use of the device and VLE may coincide with other therapies (such as physio) which may exert an influence on their motor skills abilities, and therefore their ability to control the VLE.
- Typical rooms in which the device and learning materials are used are quite warm, with adequate lighting levels. Any device should be temperature tolerant.

The analyses of the physical and technical factors that may exert an influence of the Design Specification Document have been identified through consultation and observational case study. Together they form the environmental analysis, which in turn helps to develop the Usability Study (deliverable 5) alongside the task and user analyses.

3.7 *Distillation of Design Criteria*

The distillation of design criteria to form the Design Specification Document is nearing completion. The criteria are evolved from The Parallel Products Review (reusability of existing design solutions and features), The User Team (interactive storyboards), The Usability Team (expert design storyboards), the Baseline Evaluation (identification of usability problems of existing switch devices with prototype VLE and learning software) and the Usability Study (design considerations from user, task and environment analyses).

Earlier work on the assessment of assistive/adaptive devices for people with a physical disability using VR (Brown et al, 1997) defined a list of key desirable features for any proposed input device. These key features were that any innovative device should be robust, modifiable, adjustable, reliable and affordable. This study considers the necessity of an additional feature; its 'manufacturability' (especially in terms of cost, including the expense involved in developing a completely new production line for a potentially low volume sales product). Whilst it is imperative that this new target feature does not stifle creativity and innovation, its consideration is none the less vitally important if any new assistive/adaptive device is to reach the market and hence repay the research investment (e.g., funders and user's with disabilities expectations that their research effort is not wasted).

Traxsys will provide input to the Design Specification Document to satisfy this requirement that any proposed switch-based device should have sufficient attention paid to its manufacturability.

4. INITIAL CONCLUSIONS

An ideal outcome of this research would be the identification of a design solution for as many people with physical and associated cognitive disabilities as possible. However, they represent a heterogeneous group of people (as observed in the User Analysis) and vary in abilities along many dimensions, for example, vision and hearing abilities, gross and fine motor abilities, cognitive abilities, understanding and use of language. Neither does it follow that just because someone has low ability on one of these dimensions that they hold a similar position along one of the other dimensions.

It is highly likely that there is no single functional solution for such a large heterogeneous user group, but with the possible addition of bespoke assessment software and a contemporary database of existing switch based product information, a solution can always be recommended, if not immediately provided.

The best fit solution for the majority of the target user population will be however identified in the Design Specification Document, and this represents the distillation of all of the design stages of a proven and tested methodology to develop peripheral devices for people with a disability using VLE. The imminent prototyping and re-evaluation of this device will facilitate its comparative analysis with existing solutions to identify performance gain (against an established set of usability metrics and baseline study) and design modifications before final manufacture and distribution.

A bespoke piece of software may also be developed for the purpose of assessing users' assistive/adaptive device need. This additional software solution will be used in conjunction with the assistive/adaptive switch-based device to greatly increase its potential application, and ensure all users are able to access the VLE and learning materials.

The investigation of the use of blue tooth technology, as a means to provide user recognition and automatic calibration, as well as wireless capability appears feasible. The importance of the peripheral's aesthetic design, a theme all too commonly neglected in the development of assistive/adaptive devices, has also been highlighted. The role of users in determining these design parameters is vital and a priority in design. A novel iterative and inclusive storyboarding process to facilitate these user sensitive design processes has been developed.

Acknowledgements: The authors would like to thanks all staff and partners of the Portland Partnership for their contribution to this research and its publication.

5. REFERENCES

D J Brown, S J Kerr and J Crosier (1997), Appropriate input devices for students with learning and motor skills difficulties: Report to the National Council for Educational Technology, UK.

D J Brown, H Neale, S.V Cobb and H Reynolds. (1999), The development and evaluation of the virtual city. *Int J Virtual Reality*. **4**.1, pp. 28-41.

D J Brown, H M Powell, S Battersby, J Lewis, N Shopland and M Yazdanparast (2002), Design Guidelines for interactive multimedia learning environments to promote social inclusion. *Journal of Disability and Rehabilitation*, **24**, 11-12, pp. 587-599.ISSN 0963-8288, Taylor and Francis, Editors David J. Brown and David Rose.

J D Hall (1993), Explorations of Population Expectations and Stereotypes with Relevance to Design. *undergraduate thesis, Department of Manufacturing Engineering, University of Nottingham.*

R Soos (1998) A Hybrid Wheelchair Controller for People with Physical Disabilities. *Unpublished BSc thesis, Computer Studies, Nottingham Trent University.*

T L Lannen, D J Brown and H M Powell (2002), Access to Virtual Learning Environments for People with Learning Difficulties. *Journal of Disability and Rehabilitation*, **24**, No.11-12, pp. 578-586. ISSN 0963-8288, Taylor and Francis, Editors David J. Brown and David Rose.

P J Standen, T L Lannen and DJ Brown (2002), Control of virtual environments for people with intellectual disabilities. In Keates S, Langdon P, Clarkson PJ, Robinson P (eds) *Universal access and assistive technology*, Springer-Verlag: London, pp. 63 - 72.

Interaction via motion observation

M A Foyle[1] and R J McCrindle[2]

School of Systems Engineering, University of Reading,
Reading, UK

mfoyle@iee.org, r.j.mccrindle@reading.ac.uk

www.sse.reading.ac.uk

ABSTRACT

The main method of interacting with computers and consumer electronics has changed very little in the past 20 years. This paper describes the development of an exciting and novel Human Computer Interface (HCI) that has been developed to allow people to interact with computers in a visual manner. The system uses a standard computer web camera to watch the user and respond to movements made by the user's hand. As a result, the user is able to operate the computer, play games or even move a pointer by waving their hand in front of the camera. Due to the visual tracking aspect of the system, it is potentially suitable for disabled people whose condition may restrict their ability to use a standard computer mouse. Trials of the system have produced encouraging results, showing the system to have great potential as an input medium. The paper also discusses a set of applications developed for use with the system, including a game, and the implications such a system may have if introduced into everyday life.

1. INTRODUCTION

The invention of the electronic computer in the 20th Century has probably changed the way in which we live in more ways than any other invention. Whilst it may be ironic that Thomas Watson, chairman of IBM famously said in 1943 "I think there is a world market for maybe five computers", computers are now so widespread that considering a modern society without them is almost impossible.

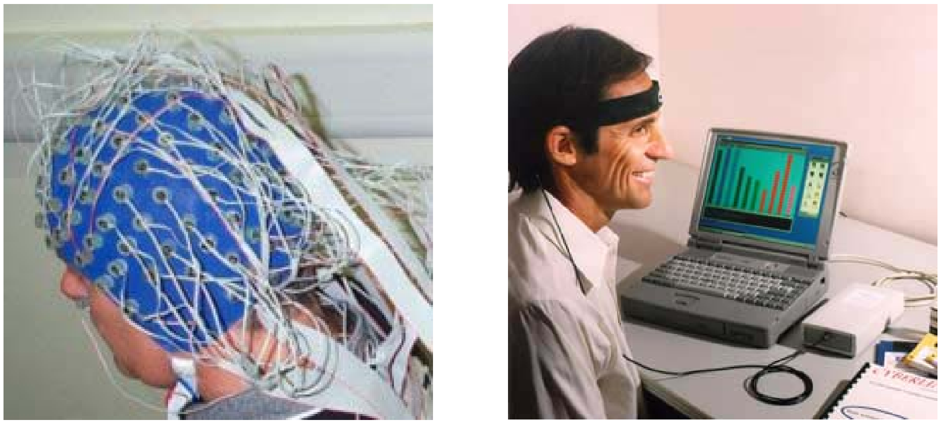
Advances in chip fabrication have seen processors become smaller and faster, storage capacities have increased phenomenally, and we have seen the computer merge with other household devices, producing so called convergence devices. Yet today, over 20 years on from the first commercial graphical computer operating systems, the principal mechanisms for data input are still the humble keyboard and mouse. In fact, the modern computer mouse stems from research developed over 40 years ago, such as the Sketchpad (Sutherland 1963) and work undertaken at Stanford (Engelbart 1967) in the late 1960s.

The aim of this project was to develop a novel and exciting method of interacting with computers, which would take advantage of these technological advancements, and which would also aid disabled people whose condition may restrict their ability to use a standard computer mouse.

The system that was produced, known as the IMO (Interaction via Motion Observation) system, is based upon a standard computer web camera, generally used for video conferencing. In conjunction with the camera, a software system was developed that is used for interpreting the images captured by the camera, so that they can be processed and turned into useful input data. The system allows the user to interact with the computer in a visual way, without needing to make physical contact with an input device, or wear any additional tracking equipment (e.g. special gloves or head gear).

2. BACKGROUND

A majority of the research being conducted into developing new input devices is being performed in order to develop systems which can been used as assistive devices for disabled persons. In some of these cases, the computer provides the user's only channel of interaction with the real world. One example of this type of technology is the Brain Computer Interface (BCI), which has been developed for use by people with severe motor disabilities. The use of electroencephalograms (EEGs) allows brain activity to be observed, and in turn the signals extracted can be used to operate a computer. The EEGs are captured from the person's brain by the attachment of an array of nodes to the outside of their head, as shown in Figure 1(a). Work carried out at the Wadsworth Center in New York has been used to drive both computers and simple prosthetic devices.



(a) *(b)*

Figure 1. *(a) A person's brain activity being measured through EEG signals, and (b) The Cyberlink Mindmouse*

In addition to current research projects there exist several commercially available devices, one example of which is the Cyberlink Mindmouse. The device consists of a headband that uses 3 sensors to detect electrical activity from the forehead and convert them into signals to drive the computer. Figure 1(b) illustrates a typical setup of the Mindmouse showing a user wearing the headband, which is in turn connected to the computer via a processing and control box. However, the problem with these devices is that they are very expensive and not readily available. For example, the Mindmouse retails at over $2000 (US), and devices using EEGs are typically experimental or hospital based.

In addition to these EEG-based interfaces, other methods of interaction are also being developed such as eye tracking. These systems typically use small cameras mounted onto the frames of spectacles, with the cameras positioned such that they can monitor the position of the user's pupil.

Our study revealed that in general the systems being developed are expensive, require special application of specific technology and generally require the user to wear some kind of monitoring device. These factors mean that devices based on these technologies are generally expensive to manufacture, and as a result are unlikely to be available to the mass market. It became clear that for any novel input device to be widely used, it would need to meet the following criteria:

- The hardware required for the device should be readily available, and relatively low cost.
- The device should not require any kind of special monitors or devices to be worn by the user.

These criteria became important influences in the development of this project, as it was felt that they were important in ensuring that the resultant input device would be both novel and accessible to all potential users.

In order to satisfy the first criterion, it was decided that using a piece of hardware which is available "off the shelf" would be desirable, particularly if the hardware already incorporates a computer connection interface. A study was conducted into the different types of computer peripherals available that might be suitable for use as a novel input device. Clearly there exist a large number of devices that are designed for user interaction in one form or another, such as gamepads and joysticks. However, these devices are related by concept to computer mice, and it was felt, do not represent a novel approach to user interaction.

The study highlighted that two "non-mechanical" devices exist which have the potential to be part of a novel input system. The first is the computer microphone, which is known for its use as a component in vocal recognition systems, and the second is the 'web camera' – a small, low resolution video camera used for video conferencing. Since voice recognition packages are commercially available, it was decided that the use of a web camera as part of an input device was both novel and interesting, since visual input systems are generally not very widespread.

3. HARDWARE

The main aim of the work was to implement the web camera as an input device in the following way. The camera is positioned such that it is able to monitor the movement of a user's hand. The user then moves their hand around in the visual area of the camera, and a continuous stream of images is captured. The software part of the system then processes the images in real time, finds the position of the user's hand and then uses this information for whatever purposed is selected – such as a custom interface, game, etc. Thus the system is able to see how the user moves, and respond accordingly, providing interaction through the observation of the user's motion. The system developed is known as Interaction via Motion Observation (IMO).

Traditionally web cameras are used for video conferencing applications, so using a web camera as a computer input device is quite a radical concept. Several choices are available in the computer web camera market, ranging from cheap, low-resolution cameras, right up to more expensive, high-resolution and high bandwidth cameras. For the purposes of this project, a mid-range Logitech Quickcam Pro 4000 camera was used (see Figure 2), costing approximately £50. This camera was chosen for several reasons – the camera is a popular and readily available camera, and the manufacturer provides a free Software Development Kit (SDK).

Figure 2. *Logitech Quickcam Pro 4000*

The camera was connected via a standard USB port to a PC, running Microsoft Windows XP. All software developed to work with the camera was designed for, and tested on the Windows XP Operating System.

4. SOFTWARE

The main task of the project was creating the system that would be able to analyse the video stream from the camera and extract the location of the user's hand, so that its position could be used to drive an interface. Whilst the system was designed primarily to be operated by the user's hand, it was decided that the image processing stage should not look specifically for a hand shaped object. Instead, it would be preferable if the movement of any hand-sized object, such as a foot, or even a small book, could operate the system. This approach would ensure that the system would be accessible and usable by a greater range of disabled people.

The detection of an object in the captured video stream was one of the most challenging parts of the project, especially due to the main issue of visual accuracy versus system response. For the system as a whole to work well, it needs to be accurate at detecting the user's hand within the captured stream of images. Many image-processing algorithms exist for this very purpose, and generally they perform a good job at finding the required object within an image. However, these algorithms can be very complex and computational (Davies 1996, Gonazalez 2002), and so may take a few seconds to process the image. Clearly this is acceptable when dealing with static images, but for a stream of images that are being used to operate a computer interface, this renders the interface unusable, regardless of the accuracy. In addition, it is important to bear in mind that the system is an input system, and so should have a low processing requirement, as it is likely to be used with other programs which may be more processor intensive. Thus it was important to ensure that the system was accurate enough to locate the user's hand, but without the lag caused by complex algorithms.

The first stage was to filter the captured images to leave behind only the objects in the scene that were of interest. Since the system is designed to be driven by a user's hand, the object that we are primarily interested in detecting is therefore a hand. However, forcing the system to look specifically for hand shaped objects could cause potential issues. For example, there is no such thing as an 'average hand', hands can vary significantly in shape and size, and if it were possible to operate the device with objects other than hands, such as feet, then the interface may be usable by a greater audience. Systems that have been

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

developed to work specifically with hands generally build a model of the human hand (Lin 2000), use a template image (Lewis 1995) or look for particular tones of skin in the image (Lenman 2002).

Since the system is not to be restricted to just hand detection, the object recognition system needed to look for general object movement. This required knowing what has changed in the image, and then determining where the object is located and its direction of movement. Early prototypes of the system used image-differencing techniques to detect motion. However, it was found that detecting the direction of the largest movements was fairly computational, and as a result the system was unresponsive and difficult to interact with.

As a result of these attempts, it was decided that a better approach was to consider what objects were 'new' to the scene, and to assume that the biggest object was the object that was to be tracked. This approach has been used in many systems, such as in hand gesture recognition (Ziemlinski 2001), where a plain background and static lighting are assumed. With this approach, when a user places his or her hand in front of the camera, the system recognises this as a new object and derives its location. The technique employed for finding motion was to take a snapshot of the background scene and to then subtract it from each image captured.

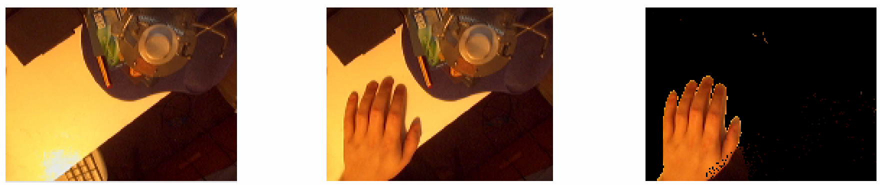


Figure 3. *A demonstration of the extraction system in operation*

This can be seen in Figure 3, which shows the background (image a), the captured image (image b) and the result of background subtraction (image c). In conjunction with a threshold technique that works relative to the intensity of the pixels in the captured image, it is possible to extract any large changes, such as hand motion. The system was tested with several different objects, including human hands, rulers and blocks of wood, and it was able to extract each of these objects relatively well. Whilst other objects, background changes and jolts to the camera could sometimes be detected in the processed image, these were "removed" by the next stage of the process, which determines the location of the user's hand (or whatever object is being used to control the system).

Once an image containing the objects has been obtained, it is then necessary to identify which of the clusters of pixels form the object of interest and the precise location of that object. Whilst techniques exist to perform this operation, due to the system response constraint, it was desirable to use a method that would work quickly, and which could be performed as the image was read from the camera. Initially, a image projection was performed, to count the frequency of pixels in each row and column. The theory behind this concept was that the object being used to control the system was generally large in comparison to any "noise" detected, and therefore the column and rows with the highest pixel counts would pass through the object. In addition, the system incorporates a weighting system, such that connected areas hold a higher value than non-connected areas, and as a result the largest object in the image will be located correctly. Tests applied during the development of the system proved this method to work successfully, although the system may be developed in the future to use more sophisticated techniques.

5. EXAMPLE APPLICATIONS

Once the basic IMO system had been created, and it was possible for the system to locate and track the movement of the user, the uses of the system in terms of computer interaction could be explored.

The first application developed was a simple program which allowed the on-screen mouse cursor to be moved around by the user moving their hand. The system performed well at this task, and it was possible to move the cursor in the same direction as the motion of the user's hands. Using the system as a replacement for the desktop mouse shows great potential, and would allow the system to be used with a large number of existing applications. Due to the difference between the resolution of the camera and the display, the motion was fairly jerky and as a result it was particularly hard to move the cursor to a precise location. However, as the specification of cameras improve and prices drop, it may be possible to improve this application in future work, by using a higher specification camera.

The second application was developed as an investigation into custom graphical input interfaces for the system. Instead of using a mouse pointer, the interface uses a concept of “zones”. Essentially the camera’s field of view is split into a grid of zones (6 for the test application described), and the system observes which zone the user’s hand is located in. The first application to use this approach split the viewable workspace of the camera into 6 zones each of an equal size. Thus, if an object was detected in the top left corner, it would be classified as zone 2. Figure 4 shows an example illustration of how the workspace is divided.

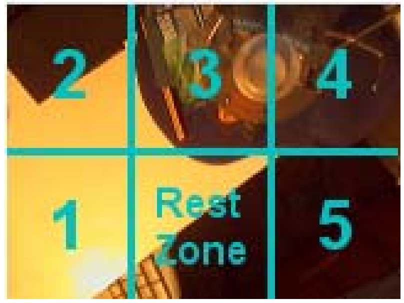
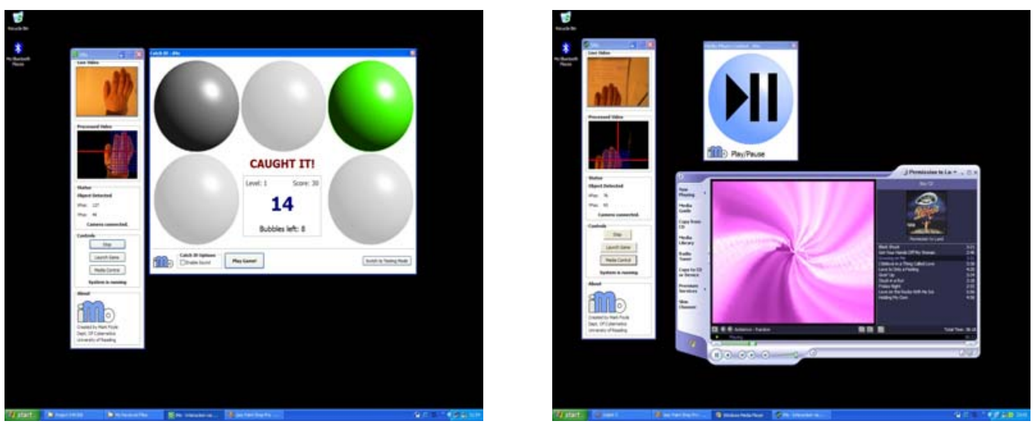


Figure 4. *How the camera workspace is divided into ‘Zones’*

An onscreen interface shows a set of objects, one for each of the camera’s visual zones. Thus if the user’s hand is located in the top right zone, this is echoed on screen by the illumination of an object in the top right area of the interface. An example of this can be shown in Figure 5(a), which shows the user selecting the top right object, by placing their hand in the top right zone. This system has the benefit that it is easier to select an object than by positioning a mouse cursor over a specific part of the screen.

Since there is no visual equivalent of a mouse button press, to signify that the user wishes to activate a command, the user simply selects an object by moving into the corresponding zone, and then pausing for half of a second. Trials of the interface showed this system to work well, and over time a user could become quite competent with the interface and selecting objects.



(a) *(b)*

Figure 5. *(a) The demonstration “Catch It!” game and (b) The Media Control program, controlling Windows Media Player*

One of the first applications developed using this interface was a reaction game, in which the computer picks at random one of the objects, and the user has to select the object as quickly as possible. Once the user has selected the object, the computer then picks another at random, and so it continues. In addition, the system records the time taken for the user to respond to the changes. An example run of the system can be seen in Figure 5(a).

In addition to the game, an application was developed which allowed interaction with a Media Player application. This application, shown in Figure 5(b), allowed the user to skip backwards and forwards through music tracks on a CD, and to also play and pause the music. The application was developed as an example of how the system could be used for a non-computing application.

6. OBSERVATIONS

Initial trialling of the IMO system indicated that it could become quite tiring to operate the device when the camera is mounted on a desk aimed horizontally at the user, as is the standard set up configuration for a web camera. This was especially true if the user was trying to reach far points of the screen. Indeed, trials of other similar systems (Lenman 2002, Freeman 1995) have also shown this to be the case.

A solution to this problem was evolved from some consideration as to how the IMO system may be used in public environments in which there is generally a lot of background motion. Due to the way in which the object recognition system was implemented (to allow users to use objects other than hands as input), any significant motion observed by the camera is recognised. Thus if there is a lot of movement in the background behind the user, for example many people walking past, then the system may become confused and recognise the motion of people walking past as the intended movement of the user. The solution to this problem was to try and cut down the background noise, preferably using some kind of static background. Since it would not be feasible or desirable to have a static background (such as a curtain) mounted behind the user, it was decided that the best approach was to rotate the camera and orient it such that it pointed down onto a surface, such as a desk. This was implemented in the system by attaching the camera to the side of the computer monitor such that it pointed downwards, with its field of view encompassing the area of desktop in front of the monitor, as shown in Figure 6.

Figure 6. *Example setup of the IMO camera*

7. TESTING AND RESULTS

After the development of the applications, a series of tests were performed to provide a measure of the effectiveness of the input system. The tests were performed using an extension to the game application, such that the response time (the time taken for the user to react and select the appropriate bubble) could be recorded. A test sequence was generated, requiring the user to select 30 random bubbles in sequence. This sequence was saved so that it could be repeated for tests with additional users. The time between the successful selection of each bubble was recorded, allowing the data to be analysed later.

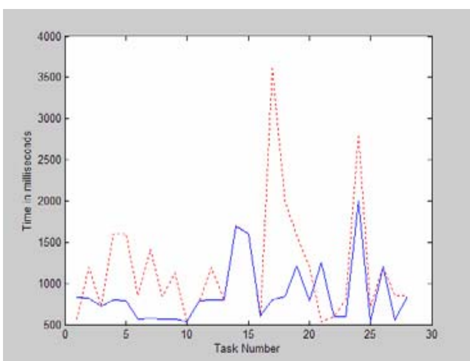
A total of 10 candidates undertook the tests, and each candidate performed the test twice, to determine whether users would become more proficient with the system as they used it more. Whilst the users were aware that they were to be tested twice, they were unaware that the second test would be identical, to ensure that they did not memorise the sequence for the second test. All times recorded were adjusted to take into account the 400 milliseconds required to trigger the target, and the different distances travelled. The final times therefore reflect the time taken by the user to respond to the change in target. Table 1 shows the average response times, per candidate, for both the first and second trials. In addition, it also shows the percentage change between the two trials.

Table 1. *Table of results*

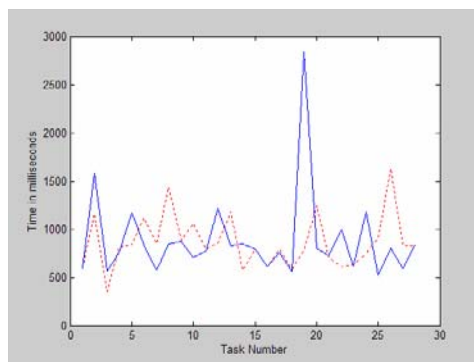
Candidate ID	First trial – average response time (ms)	Second trial – average response time (ms)	Percentage change (%)
1	863.8	885.9	-2.6
2	1187.6	845.2	28.8
3	1295.9	1766.2	-36.3
4	1139.0	1585.8	-39.2
5	1210.6	868.5	28.
6	946.3	709.2	25.1
7	1578.7	938.7	40.5
8	1244.8	950.8	23.6
9	797.7	693.6	13.0
10	1128.5	974.3	13.7

The results obtained showed that in all cases the response time was generally low (around 1 second), and that in the majority of cases, users performed better in the second trial. This would suggest that the test candidates became more competent with the system over time, and as a result were able to adjust to this method of interaction. In addition, the relatively low response time for the candidates indicates that the system responds at an acceptable speed, and is therefore suitable for this type of activity. Figure 7(a) shows the results for candidate 5 for both the first trial (dotted line) and the second trial (solid line), illustrating the improvement of the user in the second trial.

The results table shows that in three cases, the candidate's performance actually deteriorated in the second trial. Upon inspection of the results, this was due to one or two abnormal results. Figure 7(b) shows an example of one of these occurrences, taken from the results from candidate 1. It transpires that most of these abnormal results were due to loss of concentration of the candidate or a glitch in the prototype system, and thus in conclusion the system performs well as an alternative input system. If these abnormal results are ignored, then it can be seen that the performance in the second trial was either the same or better.



(a)



(b)

Figure 7. *Graphs showing the response time for each task in the 30-task trial. (a) Shows the results for candidate 5 and (b) shows the results for candidate 1. In both instances, the dotted line corresponds to the first trial, the solid line corresponds to the second trial.*

8. FURTHER WORK

As discussed previously, since the system does not require any contact forces (e.g. no pushing or pulling), it may be particularly useful for use by disabled people. If the system could be adapted such that it was built into standard household appliances, then it may be possible to make many awkward to use household appliances more accessible. As an example, the system could be embedded into current light switches, so that it would be able to turn on and off the light in a room, without having to apply the usual pressure required for a standard push button switch. This would benefit those people who may find such a procedure troublesome, due to a disability.

The prototype user interface that has been developed shows a lot of potential as a platform for a range of applications including systems that are used for serving information, such as information kiosks. Such systems generally use expensive touch-screens, but with the IMO system these could be replaced with relatively cheap cameras.

As a proposal for further study, it would be interesting to develop the system for use in Computer Supported Collaborated Work (CSCW). Users could be placed at different computers around the world, each using the IMO system, and interact within a single application. However, unlike most collaborative work systems, this would allow the users to interact through direct physical movement.

One final area for investigation is the potential application of this system in the field of rehabilitation. As the system requires physical motion for input, it could be possible to devise a virtual task, which would require the user to make a series of specific movements in order to complete the task.

9. CONCLUSIONS

This paper has discussed the development of a novel human computer interaction device, and the potential such a system has for use by disabled people, and for novel methods of computer interaction.

Since the start of the project, similar systems have started to appear in the shops for the gaming market. Systems such as the Sony EyeToy allow users to interact with basic games on a Playstation games console. Generally, these systems have only been used for gaming purposes – no mainstream system yet exists which is aimed at general computer interaction for disabled people.

The results obtained from the tests have indicated that the system performs well. In order to obtain more detailed results, we aim to develop the test applications further, and to perform more thorough testing. We also aim to test the system with people possessing limited mobility, so that the suitability of the system for disabled people can be assessed.

Whilst the system developed is primarily aimed at being used as a computer input system, there is no reason to suggest that such a system could not be used for other purposes. With the rise in convergence devices in the home, it could be possible to see such technology in consumer appliances, such as televisions and home stereos, in the very near future.

10. REFERENCES

E R Davies (1996), ‘Machine Vision: Theory, Algorithms, Practicalities’, Second Edition, Academic Press

D C Engelbart (1967), ‘X-Y Position Indicator for a display system’, USA, Reference 3,541,541, Available: http://sloan.stanford.edu/MouseSite/Archive/patent/

W T Freeman & C D Weissman (1995), Mitsubishi Electric Research Labs, ‘Television Control by Hand Gestures’, IEEE International Workshop on Automatic Face and Gesture Recognition, Zurich

R C Gonzalez & R E Woods (2002), ‘Digital Image Processing’, Second Edition, Prentice Hall

S Lenman, L Bretzner, B Eiderbäck (2002), ‘Computer Vision based recognition of Hand Gestures for Human Computer Interaction’, Kungll Tekniska Högskolan

J P Lewis (1995), ‘Fast Normalized Cross-Correlation’, Vision Interface

J Lin, Y Wu, T S Huang (2000), ‘Modelling the constraints of Human Hand Motion’, University of Illinois

I Sutherland (1963), ‘Sketchpad: A Man-machine Graphical Communications System’, Ph.D. thesis, Massachusetts Institute of Technology

R Ziemlinski & C Hynes (2001), ‘Hand Gesture Recognition’, Available from http://www.via.cornell.edu/ece547/projects/g13/exp.htm

Problems with control devices experienced by people with intellectual disabilities using virtual environments: a systematic evaluation

P J Standen[1], D J Brown[2], N Anderton[3] and S Battersby[4]

[1,3]Division of Rehabilitation & Ageing, University of Nottingham,
ADRU, B Floor, QMC, Clifton Boulevard, Nottingham NG7 2UH, UK

[2,4]School of Computing & Mathematics, Nottingham Trent University,
Newton Building, Burton Street, Nottingham NG1 4UB, UK

p.standen@nottingham.ac.uk, *david.brown@ntu.ac.uk*, *nicola.anderton@nottingham.ac.uk*,
stephen.battersby@ntu.ac.uk

www.nottingham.ac.uk/rehab, www.isrg.org.uk

ABSTRACT

Virtual environments have a role to play in facilitating the acquisition of living skills in people with intellectual disabilities, improving their cognitive skills and providing them with entertainment. However, the currently recommended devices to allow navigation in and interaction with the environments are difficult to use. Using a methodology established in an earlier study, the study aims to systematically document the performance of users with the currently recommended devices in order to i) inform the design of a usable control device or devices and ii) act as a baseline against which they can be evaluated. 40 people with severe intellectual disabilities aged between 21 and 67 years used four environments with an equal number of sessions with the different devices being evaluated. Results suggest that for navigation, the joystick is better than the keyboard but that for interaction the mouse is better than using the fire button on the joystick. Preventing slippage of the joystick base would make its use much easier and it is suggested that separate devices are retained for navigation and interaction.

1. INTRODUCTION

In the UK around 25 people in every thousand have mild or moderate intellectual disabilities and about four or five per thousand have severe intellectual disabilities (DH, 2001). These proportions are similar in many developed countries and where lower rates have been reported, for example in Taiwan, it is acknowledged that they are increasing (Lin, Wu &Yen, 2004). For the most disabled of these help will always be needed with almost every aspect of daily living yet even those who are more able will still need a degree of support to achieve the things the rest of society takes for granted. According to the 2001 White Paper (DH, 2001), people with intellectual disabilities are amongst the most socially excluded and vulnerable groups in Britain. The intention of current policy is to enable them to have as much choice and control as possible over their lives, be involved in their communities and to make a valued contribution to the world at work.

Virtual environments (VE) have a role to play in this process (Cromby, Standen & Brown, 1996). The benefits of using virtual environments (VE), for people with intellectual disabilities have been described by Cromby, Standen & Brown (1996). Initial work suggests that they are effective in facilitating the acquisition of living skills for example shopping and navigating new environments (Standen, Cromby & Brown, 1998) by children with severe intellectual disabilities. Their three-dimensional nature allows the creation of ecologically valid settings to promote activities like choice making (Standen & Ip, 2002) which people with intellectual disabilities have limited opportunity to practice (Cooper & Browder, 2001). Finally, they can provide an engaging activity for people who are frequently under occupied and denied real world opportunities (Standen, Lannen & Brown, 2002).

The work carried out so far has employed desktop VE where the environment is displayed on an ordinary computer monitor. Utilising control devices, the user's tasks are to navigate their way around the environment and interact with it. Navigation in a VE can have two meanings. Primary navigation is finding one's way in the environment. Secondary navigation is manipulating the viewpoint as seen through the

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44

computer screen. Navigation can be on a continuum between automatic, where the user is taken through the environment without any action on their part and self-controlled. In reality it is usually semi automatic, i.e. constrained by the software. So, for example, many environments employ terrain tracking where the user can only move on a horizontal plane. Within this plane there are usually 2 degrees of freedom, i.e. forward /back, turning left or right. Turning can be as for a person (pivoting on the spot) or for a car (the user must always move forward to turn). Interaction can activate objects (select item on supermarket shelf and move it into trolley), move them or cause one object to interact with another (e.g. put key in lock). These latter two functions would involve drag and drop.

A variety of devices are employed for controlling the software but Hall (1993) recommended that for navigation a joystick limited to two degrees of freedom had the greatest utility. The more functions a device possesses, the more difficult it is to operate. So, for example, when using a spaceball which has six degrees of freedom, the user with intellectual disabilities frequently became lost. Brown, Kerr & Crosier (1997), evaluating a range of affordable and robust interaction and navigation devices, also favoured use of the joystick finding it more suitable for navigation tasks than the keyboard. For interaction tasks, if drag-and-drop was not required, the touch-screen and mouse were equally effective, although the touch-screen was difficult to calibrate.

However, even the preferred devices of joystick for navigation and two button mouse for interaction are not without their problems. The two main reasons for this are the level of cognitive ability of the users and motor difficulties they experience.

Standen, Brown, Proctor & Horan (2002) tried to identify what strategies tutors employ in teaching people with intellectual disabilities to use VE designed to teach independence skills and how effective these are. Much of the time spent by the tutor in the learner's early sessions was on providing assistance with the input devices. Users experienced problems in remembering what tasks were accomplished by each device and in moving from one device to the other as many used the same (dominant) hand for both devices.

Many people with intellectual disabilities have fine motor difficulties as they suffer from conditions where damage has been caused to the central nervous system, such as cerebral palsy, multiple sclerosis, muscular dystrophy and dyspraxia. They therefore find the devices difficult to control. Lannen, Brown & Powell (2002) carried out a detailed analysis of the difficulties experienced by fourteen school aged students with intellectual disabilities using a joystick for navigation and a mouse for interaction to complete specified navigation and interaction tasks within the virtual city (Brown, Neale, Cobb & Reynolds, 1999). Many of the difficulties users experienced were due to physical ability and device construction. Finally, difficulties arose as a result of the design of the VE. With problems like these, users can become frustrated and demotivated. Using a methodology established in an earlier study, the aim of this research is to systematically document the performance of users with the currently recommended devices. This information can then be used to inform the design of a usable control device or devices and to act as a baseline against which they can be evaluated.

2. METHODS

2.1 *Design of study*

Performance data were collected on the currently used control devices which could be compared with similar data collected on any prototype developed. To reduce the effect of learning on performance, participants were allowed several practice sessions with the currently used devices before data collection started and the order in which devices were used was balanced.

2.2 *Participants*

40 people (17 men, 23 women) aged between 21 and 67 years who regularly attended a day centre for people with intellectual disabilities volunteered to take part and all met the requirement of having sufficient visual ability to see the VE on the computer monitor. Although they varied considerably in their ability, on measures of verbal and non verbal IQ, they all scored within the severely disabled range. For motor control and co-ordination, five were in the normal range, 27 showed moderate discrepancy and five showed severe discrepancy.

2.3 *Virtual environments*

Four training VE were constructed (see Figure 1), in order to evaluate the currently used devices as well as any prototypes that were developed. They were built using 3D Studio or Plasma and then imported into

Director so that all interactive elements and lighting could be coded in. In order to increase their attractiveness to users and facilitate the acquisition of navigation and interaction skills they were all designed using game format in that they consisted of varying levels of difficulty with access to each level only allowed once the correct level of performance had been achieved at the previous level. Additionally, feedback in the form of scores was available to the user. Each environment constrained different possibilities in order to test a range of uses of the control devices but without presenting the user with too many options initially. Interaction was limited to activating an object. In three of the environments the user followed an avatar. The software also collected information on task achievement (scores), time taken and collisions.

Figure 1. *Dolphins, Asteroids, Temple and Road Crossing.*

2.3.1 Dolphins: the user's view was a position behind a dolphin which they had to steer through a series of underwater hoops while avoiding natural obstacles. This required a device for up/down and left/right movement only as forward movement was provided by the software. The devices evaluated were the joystick or the arrow keys on a normal keyboard.

2.3.2 Asteroids similarly required use of a device for up/down and left/right movement only as the user's view was from a spaceship with asteroids flying towards them from different angles. Additionally interaction was required to shoot at the asteroids to destroy them before they hit the spaceship. The devices evaluated were the joystick or mouse. Each of these devices was used for both navigation and interaction i.e. when using the joystick to navigate, the fire button was used for interaction and when using the mouse the left hand button was used for interaction.

2.3.2 Temple and Road crossing required a device for movement forwards and left/right in the horizontal plane and interaction to pick up or interact with objects. The devices evaluated were the joystick or the arrow keys on a normal keyboard. Whereas for *Asteroids* the same device was used for navigation and interaction, in these environments different devices were used for interaction and navigation. The *Temple* has six levels and the user has to find an object that allows them to open a door taking them to the next level. The computer records the time taken to complete each level and the number of collisions. In *Road crossing* the user has to cross the road at a different point in each of three trials: a zebra crossing (no interaction), a pelican crossing with auditory warning and two pelican crossings with no auditory signal at a cross roads. Density of traffic was kept constant. Table 1 summarises the characteristics of each environment.

Table 1. *Characteristics of training environments*

	Forward movement provided	Interaction required	Same device for interaction	Follow avatar
Asteroids	■	■	■	
Dolphins	■		N/A	■
Temple		■		■
Road crossing		■		■

2.4 Data collection

Data collection took place in the day centre attended by the participants. Participants had sessions scheduled for once a week, which lasted a maximum of 30 minutes but could be terminated earlier if they wished. One of the researchers (NA) sat alongside them to give assistance and encouragement. The devices used were a standard 3 axis games joystick (Microsoft sidewinder, Saitek ST200 ambidextrous); standard two button mouse and keyboard. The order in which they worked through the environments was the same for each participant starting with the environment which required fewest functions from the control devices. The order in which they used the different devices varied between participants to counteract the effects of increasing familiarity on performance. Each session was recorded on videotape, with the camera positioned to view both the participant and the researcher sitting next to them. The videotapes were analysed using a method established in an earlier study (Standen et al, 2002) which yielded measures of help given by the researcher. This was described as concerning the devices (whether for navigation or interaction) or the environment and was also categorised according to whether it was verbal (information) or physical (giving assistance). Assistance is usually given where the user experiences difficulties due to the physical mismatch between their ability and the device requirements. Information is predominantly given to explain how to use the device or to negotiate the environment. The researcher also kept a diary to record any other information that might be useful but that would not be picked up by video analysis or the software data gatherer. Computer collected information on scores and collisions was adjusted for length of session. Video collected data were expressed as a percentage of session duration. Comparisons between the devices was made using the Wilcoxon test for paired data.

3. RESULTS

3.1 Games where forward movement was provided

3.1.1 Mouse versus joystick. Using the mouse allowed participants to gain significantly ($p<0.006$) higher median scores initially with the joystick and on the second (final) session ($p<0.001$) in *Asteroids* (Table 2).

By the second trial using the mouse they needed significantly ($p<0.006$) less information about navigation and less assistance when navigating on both the first ($p<0.002$) and second ($p<0.001$) trials than they did with the joystick. The amount of information given on interaction was significantly lower on the second trial with the mouse ($p<0.005$), as was the median for the tutor giving assistance with interaction ($p<0.007$).

Table 2. *Medians of performance measures for Asteroids*

	Joystick		Mouse	
Trial	1	2	1	2
Score (adjusted)	4.24	4.05	7.09	12.23
Information on navigation	13.07	6.81	9.24	0.45
Tutor assists navigation	25.96	10.25	4.46	0
Information on interaction	11.82	2.08	6.06	0
Assistance with interaction	2.34	0	1.59	0

3.1.2 Keyboard versus joystick Participants gained significantly higher scores (see Table 3) with the joystick than with the keyboard on all trials (first: $p<0.001$; second: $p<0.001$; third: $p<0.005$). There were no differences in collisions but with both devices these became more frequent with increasing exposure. No significant differences were found between the devices on any of the data obtained from the video recording.

Table 3. *Medians of performance measures for Dolphins*

Trial	1		2		3	
Device	Keyboard	Joystick	Keyboard	Joystick	Keyboard	Joystick
Score (adjusted)	1.42	2.97	2.05	2.72	2.00	2.75
Collisions (adj)	15.13	16.27	16. 71	15.59	21.55	21.52

3.2. Games where user has to effect forward movement with the navigation device.

Participants were faster with the joystick than the keyboard on each level of *Temple* (see Table 4) for which data were available and on the zebra and pelican crossings in *Road Crossing* (see Table 5). This difference was significant for level 1 ($p<0.05$) and level 5 ($p<0.02$) of *Temple* and for the zebra ($p<0.02$) and pelican crossing ($p<0.01$). There was no difference between the devices in the rate of collisions.

In terms of help given by the tutor, for *Road crossing* participants received much less information about navigation with the joystick for all three parts of the game but this was only significant ($p<0.005$) for the pelican crossing (see Table 6). When playing *Temple* they received the same amount of information about navigation with both devices. In this game more assistance was given with navigation when the joystick was being used but this difference was only significant on the last three trials (trial 4: $p< 0.035$; trial 5: $p<0.007$; trial 6 $p<0.01$) (see Table 7). In *Road crossing*, they received less assistance with the joystick but this was not significant (see Table 6).

Table 4. *Median time in minutes to complete first five trials of Temple*

Trials	1	2	3	4	5
Keyboard	4.11	4.08	3.43	3.19	3.54
Joystick	3.24	3.31	2.29	3.03	2.85

Table 5. *Median times in minutes for Road crossing*

Device	Zebra	Pelican	Crossroads
Keyboard	0.54	0.54	0.58
Joystick	0.45	0.39	0.58

Table 6. *Median percentage of time in which help was given with Road crossing*

Help given	Navigation				Interaction			
	Information		Assistance		Information		Assistance	
Device	Keyboard	Joystick	Keyboard	Joystick	Keyboard	Joystick	Keyboard	Joystick
Zebra	35.42	29.76	20.34	18.14	N/A		N/A	
Pelican	11.43	0	2.86	0	26.19	9.09	5.13	3.23
X roads	0	0	0	0	3.64	0	6.02	6.90

Table 7 *Median percentage of time in which assistance with navigation was given with Temple*

	Trial					
	1	2	3	4	5	6
Keyboard	8.82	8.67	6.63	3.56	2.8	3.92
Joystick	15.34	14.73	11.76	10.55	10.72	8.84

A comparison of the time spent receiving help with interaction suggests that less information about interaction is required with the joystick and this was significant for pelican crossing ($p<0.001$). Similar results were found for *Temple* (see Table 8) where the joystick required significantly less information about interaction on first trial ($p<0.001$), second trial ($p<0.05$) and sixth trial ($p<0.033$). Assistance with interaction was given significantly less for the joystick on the pelican crossing ($p<0.042$) and on the first ($p<0.05$) level of the *Temple*.

Table 8 *Median percentage of time in which information was given on interaction for Temple*

	Trial					
	1	2	3	4	5	6
Keyboard	8.14	0.98	0	0	0.7	0
Joystick	4.19	0.56	0	0	0.26	0

3.3 Additional information

For all four games, slippage of the base of the joystick was a problem for 25 to 37% of participants. Even on the last trial of the *Temple* when participants had considerable exposure the joystick was slipping for 13 out of 30 people for whom this information was available. The number of people using too much force with the joystick increased from 5 on the first trial to 13 on the last trial.

4. DISCUSSION

When navigation in the virtual environment is in the vertical plane only, as for *Asteroids*, use of the mouse allows better performance both initially and after exposure. Participants using it also needed less information and assistance with both navigation and interaction when using the mouse. As use of the joystick could be avoided by writing forward movement into the software this option should be explored further. The joystick was not at a disadvantage when compared with the arrows on the keyboard as the joystick enabled participants to gain consistently higher scores than they did with the keyboard.

When the user had to initiate forward movement with the navigation device, the joystick permitted better performance than did the arrows on the keyboard in terms of speed of achieving tasks. Participants required less information about using it to navigate which suggests that they are more likely to understand its function. There is a tendency for them to need more assistance using it and frequent occurrences of the joystick slipping were recorded. If it slips, the researcher or tutor has to hold the base steady and the lack of feedback from the device makes it less likely that the user will learn the appropriate amount of force to exert. When using the joystick, participants required less information and assistance with interaction even though for these environments a separate device was being used for interaction. This may be because if the navigation device is easier to understand and use, the user has more cognitive capacity available to handle the challenge of interaction.

These advantages of the joystick suggest that the most promising design solution would be to modify the basic games joystick to avoid some of the obvious difficulties that participants experienced, for example the slipping of the device over the surface on which it stood. It is more difficult to decide whether the joystick should combine both functions of navigation and interaction or whether separate devoices be retained for each function. Lannen (2002) evaluated a prototype two handed device that combined both navigation and interaction but found that young people with intellectual disabilities were confused about which action caused interaction, which navigation. The present study found that the mouse was easy to understand and use for interaction supporting Lannen's (2002) findings that for interaction the mouse was quicker to use than a button on her prototype two handed device. This suggests that the functions of navigation and interaction should be provided by two separate devices.

In summary, when navigation is required in a three dimensional environment the joystick once mastered does allow even some of the most disabled users to achieve better performance. These results suggest that resolving some of the physical difficulties with the joystick may reduce the likelihood of demotivation on initial usage and also allow better performance once use of the device has been mastered.

Acknowledgements: This research was carried out with the support of EPSRC award number GR/R21851

8. REFERENCES

D J Brown, S J Kerr and J Crosier (1997), *Appropriate input devices for students with learning and motor skills difficulties*: Report to the National Council for Educational Technology, UK.

D J Brown, H Neale, S V Cobb and H Reynolds (1999), The development and evaluation of the virtual city, *International Journal of Virtual Reality,* **4**, *1*, pp. 28-41.

K J Cooper and D M Browder (2001), Preparing staff to enhance active participation of adults with severe disabilities by offering choice and prompting performance during a community purchasing activity, *Research in Developmental Disabilities,* **22**, pp.1-20.

J J Cromby, P J Standen, and D J Brown (1996), The potentials of virtual environments in the education and training of people with learning disabilities. *Journal of Intellectual Disability Research* **40** *6,* pp. 489-501.

J D Hall (1993), Explorations of Population Expectations and Stereotypes with Relevance to Design undergraduate thesis, Department of Manufacturing Engineering, University of Nottingham.

T J Lannen (2002), A multi-disciplinary approach to the control of virtual environments for young people with moderate to severe learning difficulties. PhD thesis, Nottingham Trent University.

T J Lannen, D J Brown, and H Powell (2002), Control of virtual environments for young people with learning difficulties. *Disability and Rehabilitation* **24,** *11-12,* pp. 578-586.

P J Standen, J J Cromby and D J Brown (1998), Playing for real. *Mental Health Care*, **1,** pp.412-415.

P J Standen and W M D Ip (2002), An evaluation of the use of virtual environments in improving choice reaction time in people with severe intellectual disabilities. *Proceedings of the 4th International Conference on Virtual Reality and Associated Technologies*, (Sharkey, Sik Lányi & Standen, Eds) Veszprém, Hungary, pp.19-24.

P J Standen, T L Lannen and D J Brown (2002), Control of virtual environments for people with intellectual disabilities. *Universal access and assistive technology*. In (S Keates, P Langdon, P J Clarkson and P Robinson Eds) Springer-Verlag: London pp. 63 – 72.

P J Standen, D J Brown, T Proctor and M Horan (2002), How tutors assist adults with learning disabilities to use virtual environments *Disability and Rehabilitation*, **24,** *11-12,* pp. 570-577.

ICDVRAT 2004

Session X. Virtual Environments for Assessment

Chair: David Brown

Using physiological measures for emotional assessment: a computer-aided tool for cognitive and behavioural therapy

B Herbelin[1], P Benzaki[2], F Riquier[3], O Renault[4] and D Thalmann[5]

[1,2,4,5] VRLab, Swiss Federal Institute of Technology (EPFL),
1015 Lausanne, SWITZERLAND

[3]STAH, Adult Psychiatry University Department (DUPA),
1008 Prilly, SWITZERLAND

[1]bruno.herbelin@epfl.ch, [2]patrick.benzaki@wanadoo.fr, [3]francoise.riquier@inst.hospvd.ch,
[4]olivier.renault@epfl.ch, [5]daniel.thalmann@epfl.ch

[1,2,4,5] vrlab.epfl.ch, [3]www.chuv.ch/psy/dupa.htm

ABSTRACT

In the context of Cognitive and Behavioural Therapies, the use of immersion technologies to replace classical exposure could improve the therapeutic process. As it is necessary to validate the efficiency of such a technique, both therapists and VR specialists need tools to monitor the impact of Virtual Reality Exposure on the patients. According to previous observations and experiments, it appears that an automatic evaluation of the Arousal and Valence components of affective reactions can provide significant information. The present study investigates a possible solution of Arousal and Valence computation from physiological measurements. Results show that the dimensional reduction is not statistically meaningful, but the correlations found encourage the investigation of this approach as a complement to cognitive and behavioural study of the patient.

1. INTRODUCTION

Cognitive and Behavioural Therapies (CBT) tend to help patients to cope with their anxiety (or phobias) by successive imaginary, mediated, or in-vivo exposures. The efficiency of this form of psychotherapy has been recognized since the early eighties. More recently, with the development of the immersion technologies, a variant of exposure has been developed; the Virtual Reality Exposure (VRE). Because this technique presents several advantages (flexibility, cost), researchers aim to determine if this kind of artificially mediated stimulation has the same impact on the patient as the classical in-vivo or imaginary exposures.

One approach to determine the efficiency of VRE is to develop these experiments on large cohorts and on different phobias in order to compare the results with the classical techniques; this would be a long process. However, by taking into account the actual knowledge in Virtual Reality and Affective Computing, we can already provide some strong theoretical arguments in favour VRE. The key factor is certainly the Sense of Presence (SoP); if we had means of proving that the SoP during a VRE is comparable with the one during an in-vivo exposure, then the efficiency of both methods would be comparable as well.

The problem when studying the Sense of Presence is to find a valuable means of evaluation. In effect, trying to give a single value to the SoP may even be a non-sense. What is concretely done is the analysis of quantitative and qualitative indices of the subject's reactions during an immersion experience: cognitive response, overt behaviour and emotional states. This is also what psychologists do when they evaluate their patients all along the therapeutic process. Actually, the need is not to replace them, but to provide them with an automatic 'sensor' of the patient's state. Toward that end, studies on Affective Computing propose several means and techniques to perform a transformation of human factors into computable data. Among them, the Arousal/Valence model retained our attention: it covers a large spectrum of emotions, it is widely used in psychology, and it has already been related to the Sense of Presence.

Our working hypothesis for studying the efficiency of VRE is the following; the observation of the patient's reactions with Arousal and Valence indices only can provide enough information on his sensitivity to the virtual content. Aside from the possible discussions on this hypothesis, the objective of this paper is to show that we can obtain a correct evaluation of Arousal and Valence with physiological measures

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

exclusively. In order to build this tool, we have to deal with several constraints: dependency on an individual, difficulty to induce emotions, choice of physiological signals, selection of computational models, evaluation of the reliability, etc. As a first approach, we designed an emotion induction protocol involving one actor over several sessions during which we measured five different physiological signals. Statistical analysis has then been performed to correlate these data with the emotional classes, the cognitive evaluation of Arousal and Valence, and their expected values.

2. OVERVIEW AND OVERLAYS OF CBT AND VR

2.1 Sense of Presence in Virtual Reality Exposure

Just as Human-Computer Interaction, VR is a domain of computer science which is highly dependant on the understanding of human behaviours. The Sense of Presence during an immersion experience is commonly defined as the sense of 'being there' (Slater 1994) or as the 'illusion of non-mediation' (Lombard and Ditton 1997). Regardless of its definition, Presence is generally evaluated with questionnaires. However, Usoh et al (2000) have shown that there is no significant difference between the answers to Presence questionnaires of subjects having a real experience and those having a virtual one. Moreover, in the context of a therapy using VRE, the answers to a questionnaire cannot distinguish the part related to the patient's troubles to the one related to his presence in the virtual environment. For instance, the experiments conducted by Pertaub et al (2002) have shown that, for social phobic subjects, the feeling of being present in front of a virtual assembly was highly influenced by the attitude of the virtual actors. As a result, the comparison between the SoP during in-vivo and virtual exposures cannot be done in this classical way.

On the other hand, observations of the subject's overt behaviours, although very useful for therapists, are hard to conduct and to quantify. In practice, only monitoring the performance in the achievement of a task can provide numerical estimations (e.g. navigation and orientation in space as in (Prothero et al, 1995)). But this will not exactly indicate how much the subject was impressed by his Presence in the virtual environment.

Finally, despite the difficulty to interpret the data, physiological measurements have at least the advantage of being universal and objective. Dillon et al (2000) already proposed to use the physiological measurement of Arousal to indicate the presence during immersion. Wiederhold et al (2003) also concluded that "percentage change in heart rate and skin resistance had a high level of correlation with Presence, degree of realism, and immersiveness". The present study does not pretend to directly link physiological measures to the Sense of Presence, but these arguments encouraged our investigations about the evaluation of low-level human reactions.

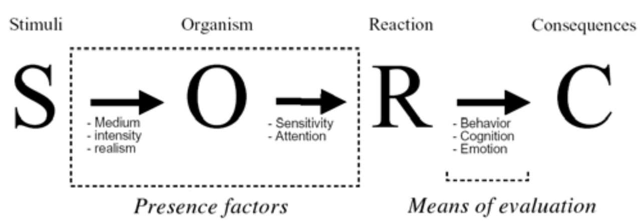


Figure 1. *The S-O-R chain of human behaviour*

2.2 Previous Experiments in VRE

Our collaboration between therapists and computer scientists started some years ago with the ambition of developing an immersive platform for the treatment of Social Anxiety Disorder. In a preliminary study (Herbelin, Riquier et al, 2002), we started to evaluate the stress generated by the exposure to a virtual assembly made of eyes looking at the subject. The subjective stress evaluation (cognitive) and the physiological measures (pulse and Electro-Dermal Activity) appeared to be in strong correlation with the expected reaction previously computed on the Liebowitz scale (a self-assessment questionnaire to appreciate the degree of anxiety of the subject in many usual social situations). Afterwards, we studied in more details the interconnections between the stress generated and the Sense of Presence during this immersion (Herbelin et al, 2002). According to the S-O-R model of human behaviour (figure 1) commonly accepted by the psychiatric community, we started to emphasise the importance of the observation of the subjects' reactions for both domains and pointed out the advantages of developing emotional assessment tools:

- Provide therapists with an easy-to-use computer-aided tool: following the patients' emotional reactions during the exposure sessions can improve the management of the therapy efficiency.

- Give to VR specialists a way to compare the efficiency of various immersion protocols: thanks to emotional assessment, the SoP could be confirmed or denied objectively.

Figure 2. *Social Anxiety exposure environment for public speaking training: the virtual environment (left) and the wide-screen immersion (right).*

We recently developed a realistic simulation to train social phobic students to present their oral exams: a virtual assembly with typical listeners' attitudes was shown on a wide screen (figure 2) or into a head-mounted display. During the therapy of two students, we focused on the patients' behaviours. Special attention was paid to the gaze of the subjects who were trained to observe the audience and to hold the gaze of the virtual humans. The support of VRE was globally considered as beneficial to the therapy (i.e. one passed his exams), though from a scientific point of view, we didn't obtain enough valuable elements to comfort this conclusion and to analyse more precisely the patients' reactions. This lead us to study the possibilities offered by affective computing systems.

2.3 Affective Computing

Picard (1997) has defined affective computing as "computing that relates to, arises from, or deliberately influence emotions". This covers the examination of media content – the stimuli – as well as the analysis of affective states – the reaction. The fully computational extraction of affective content of videos made by Hanjalic and Xu (2001) illustrates perfectly the first case. In the second case, the use of physiological measurements is often chosen to represent internal affective states. For instance, Wang et al (2004) correlate the galvanic skin response to the intensity of the emotion. Moreover, the difference in the objective notwithstanding, affective states could be represented on an Arousal/Valence graph in both cases.

The simplicity if this model is indeed very attractive. As shown in figure 3, a large scope of emotions can be labelled with only a Valence (unpleasant or "negative" to pleasant or "positive") and an Arousal (drowsy or peaceful to exited or alert). However, this apparent simplicity is extremely subjective to human beings, and the computation of those indices hides a great complexity that Hanjalic and Wang overcome by selecting 'arbitrary' digital features. To avoid this, other researchers (Healey and Picard 1998; Rani et al 2003) designed protocols to learn and optimise the correlations between physiological data and affective states. Whatever the algorithm (statistical or fuzzy respectively), the principle is the same: record various physiological signals, compute several parameters, and operate the classification.

3. METHODOLOGY AND EXPERIMENTS

3.1 Principle

If isolated from the CBT-VRE context, the experiments we conducted simply consisted in recording the physiological signals on a person trying to self-induce five classes of emotions situated at the extremes of the theoretical Arousal-Valence model. In addition, the subject estimated his subjective Arousal and Valence each time. We then tried to find the best correlations between several features derived from the physiological measurements and the theoretical/estimated Arousal and Valence.

3.2 Protocol

There is no ideal way to induce specific emotions and to ensure the person has felt 'exactly' the expected emotion. Anyway, one of the most widely used methods to induce an experimental mood is the Velten (1968) Mood Induction Procedure (the subject is instructed to try to feel the mood expressed in a card). In a similar way, and in order to gain advantage from the ability of professional actors to deeply feel the emotions they are playing, Healey and Picard (1998) asked an actress to induce an emotion by self motivation only. We did the same and asked a professional actor to perform all the experiments with self induction of emotions.

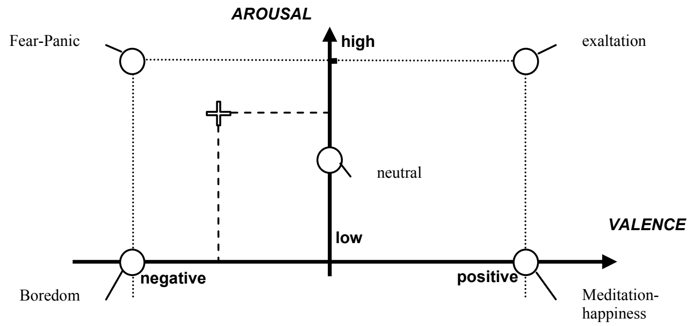


Figure 3. *The affective state in the 2D space Arousal/Valence*

Moreover, by referring to a single person, we "maximize the chances of getting a consistent interpretation for each emotion" (Picard 2001).

We used the Physio-Recorder™ from Vienna Test System Corp. to measure five physiological signals: skin conductivity level (the Electro-Dermal Activity is given in micro-Siemens), frontal *electromyography* (*venter frontalis* EMG, in micro-Volts), skin temperature (Celsius degrees), breathing frequency (abdominal and thoracic amplitude together, in cm), and pulse frequency (Heart Rate, measured with a photoplethysmo sensor strapped to a finger, expressed in beats-per-minute).

Each session starts with a relaxation phase requested by the actor and for sensors to reach stable assessment values (e.g. 10 minutes for surface temperature). Then, still lying down, the actor decides to start the inductions phases by acting on a guiding software: this little application displays the current emotion to induce and times two minutes before beeping. Then, a subjective Arousal and Valence evaluation screen appears and proposes to select values on a 1-9 scale associated with the Self Assessment Manikin from Bradley and Lang (1994) (see figure 4). One session is composed of five major emotions: "NO EMOTION / NEUTRAL", "FEAR / PANIC", "BOREDOM", "HAPPINESS / MEDITATION", and "EXHALTATION". For each of those, the timing and the subjective Arousal and Valence values are stored.

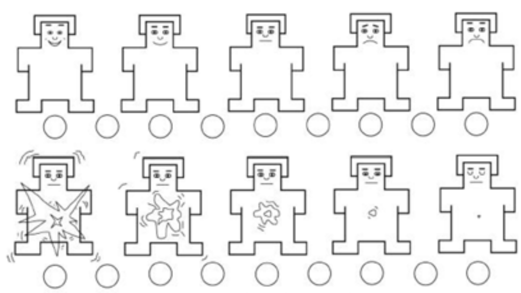


Figure 4. *Iconographic SAM rating: Valence (up) and Arousal (down) (Morris 1995).*

3.3 Data Collection and Features Extraction

During the four months of experiments, we regularly made recording sessions of approximately thirty minutes (including the relaxation phase). Due to sensor failures, the first recordings were considered as a training period for the actor, and we finally obtained ten full sessions.

The database of physiological signals is chopped according to the time sheet and regrouped by emotions (ten times five records of two minutes each). For each record X we compute the six parameters proposed by Picard (2001) on the N values (120 seconds at 10Hz gives N=1200): the mean of the raw signals (Eq.1), the standard deviation of the raw signals (Eq.2), the mean of the absolute values of the first differences of the raw signals (Eq.3), the mean of the absolute values of the first differences of the normalized signals (Eq.4), the mean of the absolute values of the second differences of the raw signals (Eq.5) and the mean of the absolute values of the second differences of the normalized signals (Eq.6).

$$m_X = \frac{1}{N}\sum_{n=1}^{N} X_n \tag{1}$$

$$s_X = \sqrt{\frac{1}{(N-1)}\sum_{n=1}^{N}(X_n - m_X)^2} \tag{2}$$

$$d_X = \frac{1}{(N-1)} \sum_{n=1}^{N-1} |X_{n+1} - X_n| \tag{3}$$

$$\tilde{d}_X = \frac{1}{(N-1)} \sum_{n=1}^{N-1} |\tilde{X}_{n+1} - \tilde{X}_n| = \frac{d_X}{s_X} \tag{4}$$

$$e_X = \frac{1}{(N-2)} \sum_{n=1}^{N-2} |X_{n+2} - X_n| \tag{5}$$

$$\tilde{e}_X = \frac{1}{(N-2)} \sum_{n=1}^{N-2} |\tilde{X}_{n+2} - \tilde{X}_n| = \frac{e_X}{s_X} \tag{6}$$

3.4 Feature selection and evaluation

Among the 30 features collected (6 for 5 signals), we need to extract the most representative. We conducted two approaches in parallel. The first consists in determining if computed features can be correlated to the Arousal and Valence as suggested in literature. A simple Pearson correlation was computed between features and self-measured Arousal/Valence over the 50 records in order to identify the most representatives.

The second approach consists in computing the Discriminant Function Analysis as commonly done when dealing with physiological signals (Nasoz et al, 2003; Christie et al, 2004). Since we want to isolate the two that fit best to the Arousal and Valence, we chose the two-group discriminant analysis, also called *Fisher linear discriminant analysis* (as in Vyzas (1999)). It allows for reducing dimensionality by finding a linear projection of the entry data to a space of fewer dimensions where the classes are hopefully well separated. According to the dimensions of our learning feature matrix (fewer training points per class than the number of total features; matrix is rank deficient), we had to apply a variation of the traditional Fisher projection algorithm by first projecting the data matrix into an ortho-normal basis [N x N] (where N is the number of training points) and produce a matrix of full rank. We then tested this projection considering three kinds of class label: 5-class-emotion, 3-class-arousal and 3-class valence. Then we focused on a resulting discrimination of the physiological features mapped on Arousal and Valence. Evaluation of the discrimination has been carried out with K-nearest-neighbours classification algorithm and the leave-one-out method has been chosen for cross validation. Here is the simplified procedure applied to each data point: (i) the data point to be classified is excluded from the original data set and the remaining data considered as the training set, (ii) the subsequent fisher projection matrix is computed from the training set and both training data and testing point are projected down to the two best eigenvectors of the Fisher projection matrix, (iii) the data point is then classified according to the KNN principle based on the Euclidian common measure distance, (iv) finally confusion matrices are calculated for the various classifications considered.

4. RESULTS

4.1 Combination of best features

The three best correlations between features and both arousal and valence were chosen and combined linearly (to keep the sign of the correlation) to compute potential estimators for Arousal (Eq.7) and Valence (Eq.8).

$$A = \frac{\alpha_A.\tilde{d}_{SCL} + \beta_A.\tilde{e}_{SCL} + \chi_A.s_{PULSE}}{\alpha_A + \beta_A + \chi_A} \qquad \text{where} \qquad \alpha_A < 0, \beta_A < 0, \chi_A > 0 \tag{7}$$

$$V = \frac{\alpha_V.\tilde{e}_{PUA} + \beta_V.d_{PULSE} + \chi_V.e_{RESP}}{\alpha_V + \beta_V + \chi_V} \qquad \text{where} \qquad \alpha_V < 0, \beta_V > 0, \chi_V > 0 \tag{8}$$

Concerning the estimation of the intensity of the emotion, our results are compliant with numerous studies that suggest that SCL and ECG signals are the most involved (indeed, correlations are around 50%). We suggest that the standard deviation of the heart rate is significant for the evaluation of the intensity of emotion (i.e. if your heart beat faster than usual, you are in a more intense affect), as are both degrees (n+1 and n+2) of variations of the skin conductivity level (i.e. Arousal is inversely proportional to the decrease of stress). Concerning the estimation of the valence of the emotion, very few papers propose an estimator. Our best correlations were only around 25% but were coherent with what Simons et al (1999) suggest: "The relationship between stimulus valence and heart rate was linear […] with the greatest deceleration associated with negative images and the least with the positive images". We also raised the probable influence of the Valence on the breathing variations. An example of A/V plot is given in figure 5: hopefully, the differences in Arousal are significant and we can distinguish positive and negative emotions. However, we observed various inconsistencies in the other sessions: an emotion is misplaced, the linear order of Valence is not respected, etc. This clearly shows the limitation of this approach and confirms the weakness of the correlations.

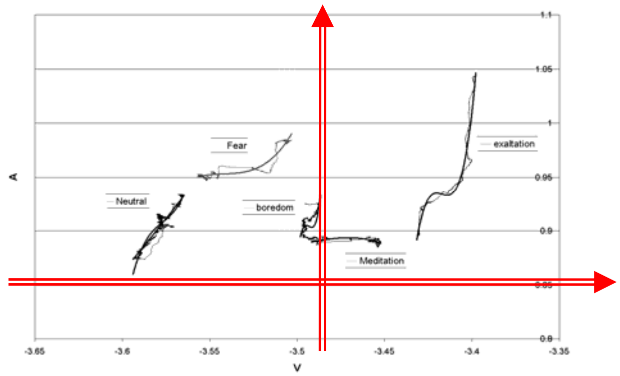
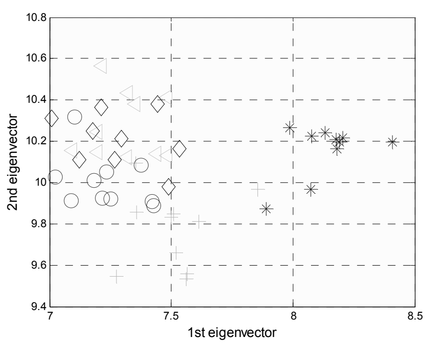


Figure 5*. Affect curves plotted on the V/A space: a possible interpretation of these factors according to the Arousal and Valence theoretical axis (with* $\alpha_A = -2, \beta_A = -2, \chi_A 0.5$, $\alpha_V = -3, \beta_V = 3, \chi_V = 1$).

4.2 Classifications

Results of the various classification protocols have very low classification rates. For example, when considering the five emotional classes (figure 6), only 24% of all the decisions of the algorithm lead to their original label. Neutral and Fear are best identified by our algorithm, although with weak success rates respectively 60% and 30%. Other classes are largely less than a random guess.



		KNN classification					
		Neutral	Fear	Boredom	Meditation	Exaltation	Total
original class	Neutral	**6**	2	0	1	1	**10**
	Fear	1	**3**	1	1	4	**10**
	Boredom	1	2	**2**	1	4	**10**
	Meditation	2	1	5	**1**	1	**10**
	Exhaltation	2	3	1	4	**0**	**10**
	Total	**12**	**11**	**9**	**8**	**10**	**50**

		KNN classification (%)					
		Neutral	Fear	Boredom	Meditation	Exhaltation	Total
original class	Neutral	**60%**	20%	0%	10%	10%	**100%**
	Fear	10%	**30%**	10%	10%	40%	**100%**
	Boredom	10%	20%	**20%**	10%	40%	**100%**
	Meditation	20%	10%	50%	**10%**	10%	**100%**
	Exhaltation	20%	30%	10%	40%	**0%**	**100%**

Figure 6. *Fisher discrimination of 5 emotions and the confusion matrix (boredom [circle], meditation[triangle], fear [cross], neutral [star] and exaltation [diamond]).*

We have also compared two different classifications of the learning data: self-measured Arousal and Valence and theoretical Arousal and Valence. It appeared that correlations between self-measured and theoretical Arousal is already very low (-0.1310) and better for the respective Valence correlation (0.8047). We may conclude that it was easier for the actor to assess the hedonic valence component of his emotion.

When considering Arousal and Valence successively as three modal classes (low, medium and high values), we obtained better results, than for the five emotion classes (figure 7). The very low number of data in the 'very low' and 'medium' categories of self-measured arousal make impossible any kind of interpretation of this part of the confusion matrix. Instead, the relatively high 85% of correct classification of 'very high' arousal may results from a subconscious over-evaluation of the real intensity of the felt physiological states. The valence component, as confirmed by the before mentioned correlation coefficients, seems to be better evaluated by the actor (compared to the theoretical model) and this can explains why their respective correct classification rate are very much comparable (39% and 45%).

However, as for the previous numerical approach, the classification could be used as a validation or as a visual monitor of the changes in the affective state.

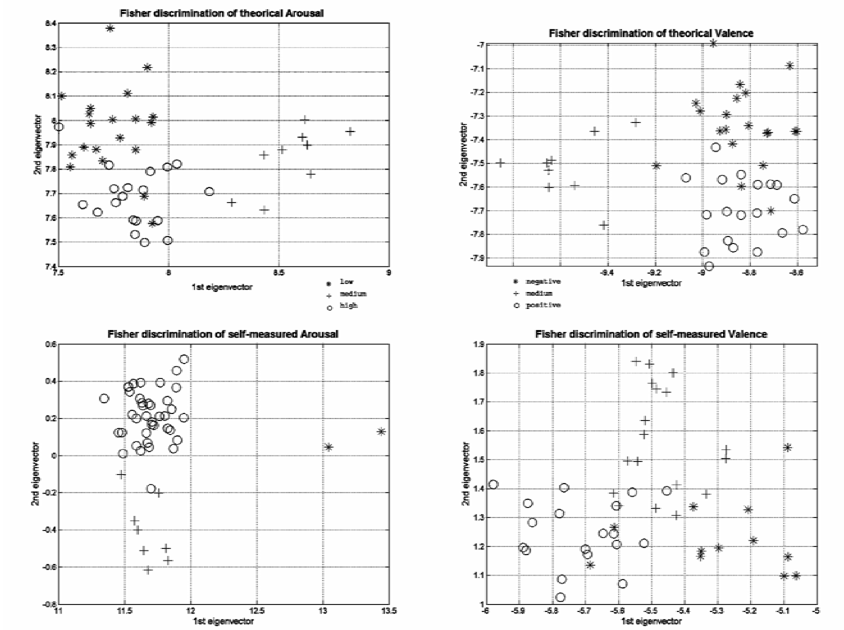


Figure 7. *Fisher's discrimination of Arousal and Valence: comparison between theoretical and self-reported classes of Arousal and Valence.*

5. DISCUSSIONS AND CONCLUSIONS

Using our feature selections and classifications, we didn't manage to extrapolate statistically meaningful Arousal and Valence from our data. First, the learning set was too small: this confirms that, in the context of therapy, patient should be studied for a very long time. Second, the projection of the physiological data into two dimensions seems too 'lossy'. As shown in the results, once the physiological features are projected on the A/V graph, it is not possible to deduce the emotional state. The classification of Arousal and Valence with two eigenvalues each would suggest that the affective state could be better represented in four dimensions.

However, additional parameters could be computed specifically for the most significant physiological signals (like the Heart Rate Variability (McCraty 2002)) and other signals could also be measured (e.g. electroencephalogram). In order to enhance the quality of the recognition, the data set has to be also trustworthy. With self-induction protocol, we can't objectively verify the affective state of the subject. The correlation with the content of media stimulation could help to establish stronger reference points.

Results are non-conclusive if the method is used to identify the emotion in an unknown context. However, the systems we built are sufficient to correlate reactions of the patient with known stimuli. They could be used to assess the patient's response to visual or auditory events during a Virtual Reality Exposure. A biofeedback loop could then be established between the patient and the virtual content. We plan to use the evaluated/recognised intensity of the emotion as a motivating factor for the behaviour of the virtual humans. In the context of social phobia therapy, the attitude of the assembly could, for example, be adapted to the stress of the subject.

Finally, we think that physiological measures have to be combined with cognitive and behavioural evaluations of the subject's reactions. Our approach is driven by the close collaboration with therapists: we try to create what they imagine as interesting tool for their work. This physiological Arousal/Valence assessment tool is an example. Another is the eye tracking system we are currently testing to have a trace of the patients' gaze avoidance behaviour. All together, they could provide therapists with the ability to monitor their patients from multiple points of view: emotional, but also cognitive and behavioural.

Acknowledgements: Authors would like to thank Stéphanie Noverraz for the proof reading of the manuscript.

6. REFERENCES

M M Bradley and P J Lang (1994), Measuring emotions: the Self-Assessment Manikin and the Semantic Differential, J. Behav. Ther. Exp. Psychiatry, **25**:49-59.

C Dillon, E Keogh, J Freeman, and J Davidoff (2000), Aroused and Immersed: The Psychophysiology of Presence, In *Proceedings of Presence 2000 - the 3rd International Workshop on Presence*, 27 March 2000, Technical University of Delft, Netherlands.

A Hanjalic and L Q Xu (2001), User-oriented Affective Video Content Analysis, In *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'01)*, Kauai, Hawaii, December 2001.

J Healey and R W Picard (1998), Digital Processing of Affective Signals, *Proceeding of the ICASSP'98,* MIT Technical Report N°444.

B Herbelin, F Vexo and D Thalmann (2002), Sense of Presence in Virtual Reality Exposures Therapy, In Proceedings of the 1st International Workshop on Virtual Reality Rehabilitation, *Lausanne, Switzerland, November 2002.*

B Herbelin, F Riquier, F Vexo and D Thalmann (2002), Virtual Reality in Cognitive Behavioral Therapy : a preliminary study on Social Anxiety Disorder, In Proceedings of the 8th International Conference on Virtual Systems and Multimedia*, Gyeongju, Korea, September 2002.*

M Lombard and T Ditton (1997), At the Heart of It All:The Concept of Presence, Journal of Computer-Mediated Communication, **3**(*2*).

R McCraty (2002), Heart Rhythm Coherence-An Emerging Area of Biofeedback, *Biofeedback*, spring 2002,

J D Morris (1995), SAM:The Self-Assessment Manikin, An Efficient Cross-Cultural Measurement of Emotional Response,(Observations), Journal of Advertising Research, November 01, 1995.

F Nasoz, C L Lisetti, K Alvarez, N Finelstein (2003), Emotional Recognition from Physiological Signals for User Modeling of Affect, In *Proceedings of the 3rd Workshop on Affective and Attitude User Modeling,* Pittsburgh, PA, USA, June 2003.

D P Pertaub, M Slater, and C Barker (2002), An experiment on public speaking anxiety in response to three different types of virtual audience, *PRESENCE - Teleoperators and Virtual Environments*, **11**(*1*):68-78.

R W Picard (1997) *Affective Computing* The MIT Press, Cambridge, MA.

R W Picard, E Vyzas and J Healey (2001), Towards Machine Emotional Intelligence: Analysis of Affective Physiological State, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(*10*).

J Prothero, D Parker, T A Furness and M Wells (1995), Towards a Robust, Quantitative Measure of Presence, In *Proceedings of Conference on Experimental Analysis and Measurement of Situational Awareness,* Datona Beach, FL, 359-366.

P Rani, N Sarkar, C Smith (2003), Affect-Sensitive Human-Robot Cooperation, Theory and Experiments, In *Proceedings of the IEEE International Conference on Robotics and Automation*, Taiwan, September 2003, 2382-238.

M Slater and M Usoh (1994), Representation systems, perceptual position, and presence in immersive virtual environments, *PRESENCE: Teleoperators and Virtual Environments*, **2**(*3*):221-233.

M Usoh, E Catena, S Arman, and M Slater, Using Presence Questionnaires in Reality, *PRESENCE: Teleoperators and Virtual Environments,* **9**:497-503.

E Velten (1968), A laboratory task for induction of mood states, *Behavior Research and Therapy,* ***6:***473-482.

E Vyzas (1999), Recognition of Emotional and Cognitive States Using Physiological Data, *Master of Science diploma project,* Department of Mechanical Engineering, MIT.

H Wang, H Prendinger and T Igarashi (2004), Communicating Emotions in Online Chat Using Physiological Sensors and Animated Text, In *Proceedings of the 1st international conference for human-computer interaction*, Vienna, Austria, 1171-1174.

B K Wiederhold, D P Jang, M Kaneda, I Cabral, Y Lurie, T May, I Y Kim, M D Wiederhold and S I Kim (2003), An investigation into physiological responses in virtual environments: an objective measurement of presence, *Towards CyberPsychology: Mind, Cognitions and Society in the Internet Age*, 175-183, Giuseppe Riva & Carlo Galimberti (Eds.), Amsterdam IOS Press.

R F Simons, B H Detenber, T M Roedema and J E Reiss (1999), Emotion-Processing in Three Systems: The Medium and the Message, Psychophysiology, **36**, 619-627.

Development of a virtual reality system to study tendency of falling among older people

L Nyberg[1], L Lundin-Olsson[1], B Sondell[1], A Backman[2], K Holmlund[2], S Eriksson[1], M Stenvall[1], E Rosendahl[1], M Maxhall[1,2] and G Bucht[1]

[1]Department of Community Medicine and Rehabilitation, Umeå University, Umeå, SWEDEN

[2]VRlab, Umeå University, Umeå, Sweden

Lars.Nyberg@Physiother.UmU.SE, lillemor.lundin.olsson@physiother.umu.se, bjorn.sondell@germed.umu.se, andersb@cs.umu.se, holmlund@hpc2n.umu.se, staffan.eriksson@germed.umu.se, michael.stenvall@germed.umu.se, erik.rosendahl@germed.umu.se, masmal@hpc2n.umu.se, gosta.bucht@germed.umu.se

[1]www.umu.se/medfak/institutioner/samh_reh_eng.html, [2]www.vrlab.umu.se, [3]www.psy.umu.se/forskning/Forskare/l_hedman

ABSTRACT

Injuries related to falls are a major threat to older persons health. A fall may not only result in an injury, but also in a decreased sense of autonomy in the persons daily life. In order to be able to prevent such falls there is a need to further understand the complex mechanisms involved in balance and walking. Here we present an immersive virtual reality system in which a person can move around, while being subjected to various events, which may influence balance and walking.

1. INTRODUCTION

Falls and fall-related injuries is a serious health problem for elderly persons (Tinetti et al 1988, Luukinen et al 1994). With age there is an increasing number of diseases present among persons in the ageing population. Neuromuscular, cognitive and vascular diseases may represent disorders contributing to poor balance and a risk of falls and injuries. Old persons also often consume numerous medications, many of which may also influence alertness, blood pressure, muscle tension, respiratory function and balance in a negative way. Here the interactions between various medication represents an important issue to consider when treating older persons with concomitant disease, bearing in mind the risk of negatively influencing motor performance with the medications. Medications which on the other hand may be vital to the individual.

Some patterns regarding risk factors for falls may be described, among these are; previous balance and gait problems, mobility factors which may include a broad range of difficulties regarding for instance transfer between furniture and wheelchair transfer or trouble bending down, visual deficits, cognitive impairments, postural hypotension, acute illness among persons in people living in institutional care, previous falls clearly increase the risk of another fall and future injuries, and the combination of several of these risk factors.

The costs, associated with fall related injuries in healthcare budgets in the industrialized world, is substantial (Zetraeus et al 1997). Costs may also involve severe social and personal suffering, however often difficult to measure, but obvious for those directly involved in the care for the affected individual.

Fall related injuries and their complications may also threat life per se in old age (Sernbo & Johnell 1993, Center et al 1999) . The presence of a fall and a fall related injury may be the fist step of a downward going spiral with poor self esteem, less mobility, loss of muscle strength and sarcopenia, increased passivity and lower independency level. In general falls occur mainly during walking (Jensen et al 2002) Falls that occur while turning may even constitute an increased hip fracture risk, compared to falls related to walking in a straight line (Cumming & Klineberg 1994). During walking there are complex mechanisms involved for balance control. There is also a considerable multisensory load upon the individual at each moment. Hence it appears that there are complex mechanisms involved that decide the outcome of a fall. The risk of falling may be regarded as dependent on the individual's motor function and control, related to the environmental demands, for the individual to handle, at any given moment. An imbalance here may leave the person at risk of fall and injury (Shumway-Cook & Woollacott 1995).

Postural control and attention are important elements to maintain balance and avoid falls. Research has shown that falls in residential care tend to occur while activities at the ward is at its greatest peak level, or while the individual is engaged in an activity or confronted with an obstacle. With age related decline in systems important for postural control, there may be competition for a limited capacity. Maintaining postural stability during frequently performed tasks, such as sitting or standing has been considered automatic, but today there is evidence that cognitive information processing is also involved. Here multisensory input during a walk change over time and there is a constant demand for corrective movements and altered strategies for maintaining balance. If the critical limits of any part or in the interactions between any of the parts in this very complex system is reached a fall may be the severe consequence. Studies have shown that the ability to correctly process multisensory input may play a role in this respect. Persons who stop walking when they are asked an everyday question are prone to falling and may even have a greater risk for future fractures (Lundin –Olsson et al 1997). Performing normally easily handled balance tasks such as standing on a foam support, may be difficult for an older person if there are simultaneous attentional demands put on the individual (Shumway-Cook et al 1997). Hence it appears that in order to take the necessary fall preventative measures one must try to understand the characteristics for each individual, and probably also undertake functional studies during walks and turns.

This complexity makes it difficult to standardize methods for testing balance and fall tendency in a everyday clinical setting, if these methods are aimed to give reproducible, comparable, measures from one time to another. Here the use of Virtual Reality may represent a new tool able to produce identical environmental scenes over and over again. Changes in weather, light and sound conditions and relations to other persons present in the real environment may be set aside. With the use of a virtual environment (VE), such conditions can be standardised and identical for a series of tests and movement measurements. The use of VE also enables the test leaders to set "critical levels" for the visual disturbances and by doing this the balance performance level for a tested individual may be estimated in the future. Here the possibility to study influences of other factors such as medications, injuries, limb restrictions and more may also be tested and evaluated within such a technical system.

Considerations of avoiding falls while using the system was discussed throughout the project. Virtual systems has been used successfully for a number of applications, such as exposure to heights, virtual spiders and hostile avatars, all used in a manner safe for the tested individual, a safety that may not have been possible in real life. Safety while testing was also an issue in the development of this virtual reality system as a tool for testing and evaluating balance and walking. Here it is possible to expose an individual to events, that would represent a dangerous element for causing an injury in real life, or something that is impossible to do in real life (for instance tilting of the VE). When using this system a trained professional accompanies the test subject at all times. Emphasis was also made during the experiments to make the subjects feel safe when exposed to various parts of the equipment at first and then for the system in general. The tested individuals was also screened for simulator sickness using the simulator sickness questionnaire SSQ (Mc Cauley & Sharkey 1992).

2. METHODS

2.1 Method

We describe the development of a system to assess how attention demanding and unexpected events influence a person's capacity to control balance and movement using a virtual environment (VE).

The technical system was divided in two parts, hardware and software. The hardware in the system consists of a V8 head mounted display (HMD) in which the computer graphics is presented in colour on a screen for each eye. There is also a tracker system (Ascension Motion Star) that provides a magnetic field and provides tracker data from sensors attached to the test person.

Two SGI computers were being used to generate computer graphics and collect data from the motion tracker system. The participant was during the trial equipped with the Ascension Motion Star tracker system, with nine motion sensors (figure1). The software consists of the image generation of the VE and the management and visualization of motion tracking data. Here data from the walks can be visualised by the representation of a skeleton on the screen with trackers trails set at desired length to enhance the visual assessment of the movements. Tracker data in 3D was also being collected during the experiments in real time. All the software used in the experiments is developed at VRlab in Performer [http://www.sgi.com/developers/devtools/apis/performer.html], and OpenGL [www.opengl.org].

This system enables immersive virtual reality where the subject may perceive as being part of the environment What this means here is that the individual, while being in the VEm can look around in any direction and still be surrounded by only computer generated images, presented for each eye in such a way that these generated images appear to be the real thing - the surrounding world.

2.2 The environment and individuals tested

The VE was an immersive 3D model of a market in the centre of Umeå city, well known to the participants who all lived in this community. The VE shows the ground materials in the city square, as well as representations of all the surrounding buildings plants and other objects. Colours are presented to look like the real downtown of the city. As the individuals tested the environment they all felt familiar to the city presented as being the city of Umeå, known to them before.

The VE is however short of other virtual persons (avatars). Connected to the VE was the control system through a TCP/IP network connection. This enabled the control system to manipulate the VE in all regards, such as the start and stop of the events.

Individuals tested; due to the nature of this study, being the first of its kind, using very advanced technical tools, we wanted to include both young and old individuals, since we had received input that the system may be too complex for older individuals. In this preliminary pilot study eight persons (age 23 to 80) participated, after being asked for voluntary consent. They all had good balance performance on traditional clinical tests, and being ADL independent, they were also independent walkers. All participants also lived in their own home in the local community.

Emphasis was made that each tested person should feel as safe as possible during the experiments. For this reason time was given for the subject to get acquainted with the equipment. By for instance being given the chance to try walking around with the sensors alone without the HMD. A test leader (trained personnel) was also walking beside the subjects at all times, ready to grasp, but never holding on to, a hip belt with safety handles. This hip belt was used at all times while being part of the walking part of the test session. This was done in order to ensure that no injuries was caused while the individuals were being tested. Some of the subjects also expressed a feeling of safety just by knowing that there was someone walking beside them while they were performing the tests. On the other hand no instructions were made as to what was going to happen regarding the disturbances (the events) in the VE. The individuals tested was also allowed to walk in their own pace. This was important since the character of this study was that of a pilot study performed to make initial observations in this recently developed system and to collect background data, and to test the systems usability.

During the experiments, each person walked on a normal floor and was visually fed the familiar environment in the HMD and was exposed to different unexpected events, such as a virtual snowfall and tilting of the VE (fig 1).

3. RESULTS

There were technical difficulties evident during the experiments. External magnetic fields were present in one building and prevented the use there. The triggering of the events at the right time was another issue. In order for the test persons to have a chance to experience them as being events and not just system errors there had to be a certain timing involved. Here previous experiences in each individual may also have played a role. The virtual snowfall is such an event that may be perceived as a malfunctioning screen or a system error to someone not used to snow.

The positive results found was that the system could be used to influence walking among the participants while data was being collected. In general there was the feeling of being safe and well taken care of during the experiments. This was especially encouraging since an effort had been made to generate such a feeling before starting the experiments. The safety issues was a part of the study protocol in this manner. Evident disturbances of balance and walking pattern such as changes in speed, stride length and balance reactions like slipping were observed (table 1).

We also found that the system could be used for both with younger and older test subjects. Here no particular difference was observed regarding handling of the equipment or tolerating the system among the participants. However on must remember that the number of subjects in this study is not large enough to provide any clear evidence. For this reason we have chosen to present data from each individual as seen in table 1.

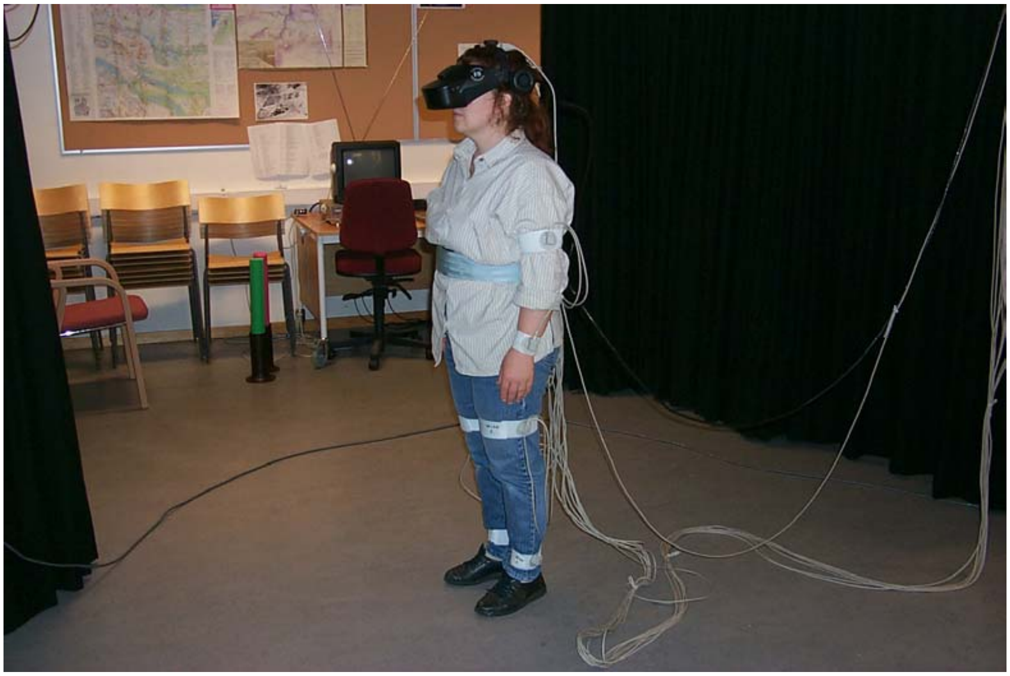


Figure 1. *Setup of the HMD and sensors.*

Two persons experienced symptoms of cyber sickness, with a SSQ score above 25 points. Measurements of walking time was being made, here walking with sensors only did not affect walking time. A striking finding on the other hand was that while walking in VE the persons generally walked more slowly. Among the events there seemed to be some difference as to the degree of impact on walking and balance. Virtual tilting of the environment had an impact on balance performance, and the system could provoke fall tendency by using this event among the test subjects. This effect was not seen in a virtual snowfall, an event that for some individuals went by without any observed impact on walking and balance. All individuals tested here had an experience of snow previously.

Table 1. *Observed balance reactions while walking in VE with and without events.*

age	80	74	72	69	34	26	23	23
sex	F	M	M	F	M	F	F	F
Walking track without events	*							
Walk 1 with events	**	*	*		*	*	*	*
Walk 2 with events	**	*	*	*			*	*

* *balance reaction*, ** *close to falling*

4. DISCUSSION AND CONCLUSIONS

This study represents our first pilot attempt to perform immersive virtual reality for the assessment of balance and walking, during events which may provoke falls. The results and conclusions must therefore be viewed with extreme caution, since the number of test persons is small and this is the first study performed. As mentioned above, evident disturbances of balance and walking pattern such as changes in speed, stride length and balance reactions like slipping were observed. We found evidence for cyber sickness, with a SSQ score above 25 points for two test subjects.

There also seemed to be differences in walking speed when the subjects were immersed in the virtual environment. Here, walking with sensors, but without the HMD did not affect walking time, but in VE the persons generally walked more slowly. Why this phenomenon occurs with slowing of the walking speed in our system still remains to be elucidated. One possible explanation is the limited field of view that the V8 HMD provides. Here the peripheral vision normally gives important cues to balance during walking that are no longer present while the person is immersed in the VE with the HMD. There may also be other technical issues involved such as how the VE corresponds to the real world during walking. It oftentimes may be so that there are imperfections as to how a subject perceives the forward motion in VE as compared to the normally experience of mowing forward when walking down the street. Maybe there is a need here to optimise the system further and evaluate the results from future studies.

Virtual tilting of the environment was the event disturbance that had the strongest impact on balance performance, and by using this visual disturbance fall tendency could be provoked. Initially the idea was that this tilting of the VE should give the feeling of actually falling while standing up. However the event rather seems to provoke fall tendency or even falls, and gives the impression that the world is leaning or tilting. The fall itself seems to generate when the tested individual tries to compensate for the abnormality generated by the computers during the tilting of the world.

The VE system was tolerated well also among older test subjects, important for its tentative use for this age group. The model needs further development, possibly by using smaller and cheaper personal computers in the nearby future, by making use of a cordless tracker system, and by finding methods to calculate and measure the observed balance reactions.

Other ways to improve the system would be by adding interactive avatars. Such interactions while measuring movements in immersive VR might add more information on the interplay of the capacity to process higher cognitive functions while maintaining a balance task, and final balance performance during such a task. By the use of different kinds of avatars and by adding sounds in the future emotion reactions may also be added as a further complication in this model.

The system clearly holds a potential for making repeated measurements in a standardized manner, a prerequisite for its clinical use, as an instrument for testing and evaluation. The system may in the future also be used as a training tool, set at different levels of difficulty, to fit individual needs. Much in the same way clinical rehabilitative training is individualized today.

Tested individuals could be provoked to having balance reactions and some also came close to falling. Still by taking certain precautions, such as careful information, and using a hip safety belt, we managed to avoid falls during the study. Still it is important to remember that the tested individuals were all independent walkers with good balance performance on clinical balance testing. When testing patients with a previous history of falls and possibly also concomitant disease even more precaution in order to avoid falls is probably needed.

Another strategy of model development is to make the VE simpler, with fever objects, but still enough to challenge and support the tested subjects. Here the development of a virtual corridor is well under way. In such a model methods to improve the measurements of each step is desired, both regarding stride length, and pace, but also regarding weight distribution and pressure. This model holds the potential for testing both expected events, that is events that allows the test person to plan ahead, and beforehand decide on a walking strategy for instance, and to test how an individual may react to unexpected events, that is an event that takes place while the test person is already performing a task with his or her movement strategy in progress. The event then calls for immediate corrective reactions in order to avoid a fall. Here the sum of multisensory input and this added level of difficulty may end up competing with limited brain and neuromuscular capacity, to maintain walking and balance. Here the virtual environment instrument may be important to find persons at risk of falling on an early stage in order to possibly undertake fall preventative measures as soon an possible.

Acknowledgements: The KK Foundation, The Knut and Alice Wallenberg Foundation, VLL, Umeå Universitet

5. REFERENCES

J R Center, T V Nguyen, D Schneider, N Sambrook and J A Eisman (1999). Mortality after all major types of osteoporotic fracture in men and women: an observational study. Lancet 353, pp. 878-882.

R G Cumming and R J Klineberg (1994). Fall frequency and characteristics and the risk of hip fractures. J Am Geriatr Soc. 42, pp 774-778.

J Jensen, L Lundin-Olsson, L Nyberg and Y Gustafson (2002). Falls among frail older people in residential care. Scand J Public Health 30(1), pp 54-61.

L Lundin-Olsson, L Nyberg and Y Gustafson (1997). “Stops walking when talking” as a predictor of falls in elderly people. Lancet. 349, p 617.

H Luukinen, K Koski, L Hiltonen and S L Kivelä (1994). Incidence rate of falls in an aged population in Northern Finland. J Clin Epidemiol. 47:, pp 843-850.

M E Mc Cauley and T J Sharkey (1992). Cybersickness: Perception of self-motion in virtual environments. Presence 1(3), pp 311-318.

I Sernbo and O Johnell (1993). Consequences of a hip fracture: a prospective study over 1 year. Osteoporosis Int 3, pp 148-153.

A Shumway-Cook and M Woollacott (1995). Motor Control. Theory and practical applications. Baltimore, Md: Williams & Wilkins.

A Shumway-Cook, M Woollacott, K A Kerns and M Baldwin (1997). The effects of two types of cognitive tasks on postural stability in older adults with and without history of falls. J Gerontol A Biol Sci Med Sci 52, pp 232-240.

M E Tinnetti, M Speechley and S F Ginter (1988). Risk factors for falls among elderly persons living in the community. N Engl J Med 319, pp 1701-1707.

N Zethraeus, L Stromberg, B Jonsson, O Svensson and G Ohlen (1997). The cost of a hip fracture. Estimates for 1,709 patients in Sweden. Acta Orthop Scand 68:13-17.

Survey of modelling approaches for medical simulators

A Al-khalifah[1] and D Roberts[2]

[1]The Department of Computer Science, the University of Reading,
Pepper lane, Whiteknights, Reading, UK

[2]The Centre for Virtual Environments, the University of Salford,
Manchester, UK

[1]a.h.al-khalifah@rdg.ac.uk, [2]d.j.roberts@salford.ac.uk

[1]www.hpivc.reading.ac.uk, [2]www.salford.ac.uk

ABSTRACT

Medical simulation, in particular that used for training and planning, has become an established application of Virtual Reality technology. This application became an active area of simulation development in many academic and commercial institutions around the globe. A reasonable number of successful commercial medical simulators have been launched, while others remain hostage in research laboratories undergoing further developments and improvements. This paper provides a dichotomy of modelling techniques in the context of deformation and cutting, giving examples of how these are applied in medical simulation, comparing their strengths and weaknesses, outlining limitations and pinpoint expectations for the future. We focus on mapping the aim of the simulator to the adoption of particular modelling approaches. A case study pays special attention to the simulation of human organs where we uncover advances and limitations in the application of these modelling approaches.

1. INTRODUCTION

Surgery is a critical process requiring both detailed planning and experienced practitioners. However, live humans are not readily available for dissection in order to support training and planning. Furthermore, the use of animals is inaccurate, expensive and introduces difficult ethical questions. A real need therefore exists for simulated planning and training. Virtual Reality technology encompasses 3D modelling, simulation and associated display and input devices that lend themselves well to the nature of surgery. Considerable research and development have been applied to this application and many tools have been evaluated, often with great success. Although early systems are gaining appreciation from some surgeons, there is still much work to do before the technology can offer the true faithfulness required for universal acceptance. Realistic results in simulating complex surgeries are still far from ideal, but the simulators produced so far are promising to be the training and simulating tools in future medical training laboratories. Due to the nature of the human anatomy and the complexity of surgical procedures, many issues must be considered while developing such systems. Factors such as, the purpose of the system and the nature of the human organs involved determine requirements on faithfulness to organ, tool, and procedure. A variety of technologies and techniques may be adapted and matched to meet these requirements. In this paper we cover these approaches in the context of deformation and cutting when applied to both endoscopic and open surgeries, and illustrate their usage by introducing a number of solutions developed in research, uncovering their advances and limitations. We emphasize that this survey only covers the main modelling approaches that are commonly used.

In section two we will introduce a number of modelling, deformation, and cutting approaches used and examples of medical simulators that have been developed around them. In section three we will make a comparison between the different techniques and highlight their advantages and disadvantages. Current limitations and future expectations of the present systems will be outlined in section four.

2. MODELLING APPROACHES

An important decision in designing a medical simulator is the choice of modelling approaches. What we mean by modelling is the way objects or models are represented, whether they are represented geometrically by mathematical procedures or abstractly by basic shapes and primitives. This decision is lead by concerns including both the body material to be simulated and the way in which it will be manipulated. A new and demanding problem might require a unique combination of approaches. Here we highlight the main approaches used together with some examples to illustrate how and why they are applied. We begin by describing modelling approaches and then explain approaches better suited to model deformation and cutting.

2.1 Object Modelling and Datasets

As datasets used in the simulated models are normally generated using different tools, and the modelled systems have distinct simulation goals to address, different modelling approaches and datasets have been used for different simulators. Here we outline the main data representation techniques and datasets used.

2.1.1 Volumetric models from scanned images. Data may be taken from Magnetic Resonance Imaging (MRI), Computer Tomography (CT), or Positron Emission Tomography (PET) scans in order to generate 3D volumetric models of the human organs. The generated models can be visualized as high quality voxel-based (3D) volumetric objects. These objects are seen as 3D array of data, where each element represents a sampled point from the acquired 3D scans. The main advantage of this method is that it maintains the original volumetric data from the scans and full volumetric representations can be produced. This technique has other advantages such as (Gibson et al, 1998),

- Voxel-based modelling is a natural way to represent 3D images, because the organization of represented data is the same as that acquired, i.e. true reflection of the original data.
- Volumetric objects can hold detailed data about the internal anatomical structure of the tissues.
- Voxel-based modelling is only suitable if the number of volumetric elements is relatively low, due to reasons mentioned below.

An example of a simulator which was developed using high resolution MRI images is the knee arthroscopy simulation system (Mitsubishi, 2001; Gibson et al, 1998). The images are first segmented, and then used to produce 3D volumetric models of the knee bones and other tissues. The simulator was developed using T-1 weighted proton density MRI images. These images then segmented into different tissue types. It was assumed that the data stored in a regular grid of evenly spaced volume elements. Good deformation results were obtained, but they were not tested against other accurate approaches such as, finite element method (Zienkiewicz and Taylor, 2000).

The main draw back of volumetric representation is that it requires high storage space and high computational power to handle such huge data. This cost of computation comes from accessing and computing the huge number of elements in the dataset. This eventually will have practical implications on the rendering speed (Gibson et al, 1998).

2.1.2 Indirect volume rendering. With this technique only the desired surfaces of the volumetric data are rendered (Radetzky and Nürnberger, 2002). This can be achieved by a process known as segmentation, which extracts the surfaces or areas of interest from the volumetric data. Segmentation subdivides the image into specific areas of distinguishing properties. The main advantage of indirect volume rendering is the increase in rendering speed obtained from the dramatic reduction in data. A further advantage is gained by the use of texture mapping, which allows the addition of realistic images obtained from real organs or structures to be applied to the surface.

To reduce the needed graphical power, ROBO-SIM developers (Radezky et al, 2000) used a combination of direct and indirect volume rendering techniques. 3D brain MRI datasets of actual patient were used to model the pre-plan procedures and to simulate views of the outer and inner surfaces of the head.

2.1.3 Surface modelling. The geometric representation of tissue may consist of surfaces as well as volumes. Surface representations are normally modelled as a set of polygons. The choice between surface and volume based models is governed by two elements: computer power and visual accuracy (Delingette, 1998). In terms of computation power, surface models are advantageous because they have less vertices/polygons to render than volume based models, but are less accurate. Furthermore, surface models may produce undesired or unrealistic deformations especially in thin representations. Surface models are good representations for modelling cavernous tissues such as, vessels and the gallbladder (Delingette, 1998).

At the University of Colorado Centre for Human Simulation, a real-time algorithm has been developed to simulate a virtual scalpel making cuts on a constructed body (Colorado, 2004). The system is able to simulate cuts using MRI data from the Visible Human database (NLM, 2003). A texture mapped polygonal representation was employed to produce the desired effects. In real-time, the texture mapped surface is updated to simulate the cutting and the resultant new surface.

2.2 Organ and tissue deformation

Deformation and elasticity are two important characteristics of the human body tissue. These unique characteristics forced the developers to use unique and challenging deformation techniques in order to model such behaviours with the required acceptable fidelity. Tradeoffs must be made between responsiveness, in terms of frame and update rate, and realism in terms of graphical and haptic detail. A general solution is probably unobtainable with current technology, but a number of approaches have been introduced to meet specific requirements. These approaches are outlined below together with real implementation examples.

2.2.1 Finite element method (FEM). This method is one of the most common approaches to model tissue deformable behaviour. It describes a shape as a set of basic geometrical elements and the model is defined by the choice of its elements, its shape function, and other global parameters (Delingette, 1998). FEM treats a problem in a continuous manner, but solved for each element in a discrete way. The problem can be solvable by adopting an interpolation algorithm within the different elements of the model (Wagner et al, 2002). The main advantage of FEM is that they can produce more physically realistic simulations compared with other approaches. Because mass and stiffness matrices remain constant over time interval and only evaluated at each time step, this produces realistic deformations. But it requires high computation, which can only be reduced if the number of nodes is reduced. Undesirably, the reduction in the number of nodes will in turn degrade the accuracy of the model.

Finite Element Method (FEM) was used to model tissue deformation for a gynaecology endoscopic simulator (Székely et al, 1999). The prime aim of the system was to produce realistic organic models with their inherited deformation characteristics to mimic the real endoscopic surgery simulation. A body in FEM was segmented into a finite number of elements. Positions and movements in the element were calculated from discrete nodal values. A discrete system of differential equations was generated for every element from a set of differential equations controlling the motion of material points of a continuum. The resultant equation then had to be integrated in relation to time. Uterus deformation has been modelled by employing a reduced volume integration approach based on absolute strain formulation, while the abdominal cavity has been modelled using rigid surfaces (Székely et al, 1999).

2.2.2 Mass spring method. Mass Spring Method is the most common approach for real-time simulations. In this technique, masses are assigned to vertices and a set of springs are allocated to connect vertices. Strut springs are also added to keep the mass-spring surface maintain certain shape and position. In real time, the deformable objects then deform in a physically-based manner after solving a mathematical problem in response to external stimuli (Bro-Nielsen et al, 1998). Mass Spring methods are easy to build and simulation levels are not as high as those for FEM, but they can produce acceptable real time simulations using today's hardware (Gibson and Mirtich, 1997). Despite these advantages, mass spring models have some drawbacks. Because they only approximate the true physics of the continuous body, spring constants are not always easy to derive and hence give inaccurate approximation of the physical behaviour of the material. Mass Spring models also have a problem called "stiffness" when the spring constants are high, this causes the system to be unstable and produces slow simulations (Gibson and Mirtich, 1997).

A standard mass-spring system was used to model surface deformation in the abdominal trauma surgery simulator (Bro-Nielsen et al, 1998). Cutting of surfaces was modelled using a number of functions applied to individual triangles. Another simple mass-spring model was used to model the arteries and other tubular models. This model was treated as the backbone of a tubular structure related to connected contours. Each one of these contours is connected to a vertex, which is then controlled by this vertex. The development of such system allowed the developers to investigate and apply different virtual technologies to model deformable objects and simulate other procedures such as, bleeding, cutting and haptic responses.

Mass Spring method was also used in the EyeSi simulator (Wagner et al, 2002). The system is a training simulator for intraocular surgery. In this system a Mass-Spring method had been used to model the following two different tissue deformation simulations:

- This simulation involves the removal of membrane. The rim of the membrane is normally connected to the retina, so in the mass spring model, the corresponding nodes are not moved. Removal or movement of the retina itself can be detected if the force on the rim nodes exceeds certain value.

- A common surgery procedure is to lift the retina by injecting salt solution beneath it. This procedure was also modelled by treating the retina as a mass-spring mesh.

2.2.3 Chainmail. This approach was first introduced by Gibson (1997). In this approach, the model maintains the original data resolution and allows the individual object elements (volumetric elements) to follow certain rules to do the required deformation. In a 3D format, each element is connected to six of its neighbours. These are, top, bottom, right, left, front, and back. When the object is manipulated, each element is tested to see if it violated certain distance thresholds it shares with its neighbours. If the distances are exceeded, the element is moved in the desired direction. If not, then the object does not move. These rules are applied to element/elements, which is/are within the influence of manipulation and then propagates to its/their neighbours. So, local deformations are generated only if distance thresholds were exceeded. Links between elements are initially set slack, so movements only occur if the distances were violated. Elasticity or deformation constrains can be modelled by setting different distance values for different object types. For example, by giving rigid objects small variations, while deformable objects are given high values.

The major advantage of this approach is that it takes benefit of all the high resolution volumetric data generated by scanners and applies simple calculations on elements to generate the required deformation results. A range of material types can be modelled using this approach, such as rigid, deformable, plastic, and elastic (Gibson, 1997). Chainmail method is also relatively easy to implement compared to other approaches.

A knee arthroscopy simulator was developed using such approach (Gibson et al, 1998). This simulator was modelled using a volumetric object representation as explained in §2.1.1.

2.2.4 3D linear elasticity. Because of the elastic and deformable nature of soft tissue, linear elasticity (Landau and Lifschitz, 1986) is also a common algorithm used to model deformations and cuts. Considering an object composed of a number of tetrahedral elements in a domain Ω, the elasticity theory problem solution can be defined by the solution of the linear system (Cotin et al, 1999)

$$\mathbf{K}\,\mathbf{u} = \mathbf{f} \tag{1}$$

Where **K** is the stiffness matrix and is symmetric, positive, definite, and sparse; **u** is the unknown displacement field and **f** is the external force. The size of this matrix **K** is $3N \times 3N$ where N is the number of mesh vertices (nodes). In general, a set of external forces are applied to the surface of the solid while some mesh nodes are fixed, otherwise a translation would occur, not a deformation. In linear theory, the behaviour of deformable model is physically correct only for small displacements (approximately 10% of the total mesh size), and less realistic for larger deformations. Another feature of linear theory is that any mesh deformation can be computed from knowing a finite set of elementary deformations (Cotin et al, 1999). Also, because most soft tissue materials behave in a non-linear fashion, linear elasticity can not always be applied.

2.3 Tissue Cutting

In order to simulate cuts in a soft tissue, there are a number of routines, which can be applied. These include collision detection, subdivision, and relaxation.

2.3.1 Collision detection. What we mean by collision within the focus of this paper is the intersection between the soft tissue and the scalpel or the cutting tool. Collision detection algorithm computes two actions, firstly, finds the intersections between the swept or cutting plane and the active intersected tetrahedral (Four triangular sided entity, which forms 3D meshes) edges, and secondly finds the intersections between the scalpel tip path and the intersected faces. Bielser and Gross (2000) introduced two types, the surface collision detection and the volume collision detection. Surface collision detection algorithm finds boundary elements that collide with the cutting tool; an axis aligned bounding volume hierarchy over the surface is normally employed. On the other hand, volume collision detection finds those tetrahedrons inside the tissue, which are split or partially cut by the tool.

2.3.2 Subdivision. Because most volumetric meshes are represented as tetrahedral, there has to be a mechanism to split or cut these tetrahedral elements. Subdivision is the main procedure to do this splitting. Bielser and Gross (2000) had used tetrahedron cut-specific subdivision patterns. Here only edges and faces that are part of the cutting face are considered. Five topological patterns were suggested. Three of them represent partial cutting, while the other two model a complete tetrahedron split.

Bielser et al (1999) have used the same five possible topological representations to cut a tetrahedron. For each of the five cases, a set of actions needed to produce the new mesh are stored in a look-up table. Insertion of new mass-nodes and assignment of connectivity are some of the actions stored in the look-up table. By mirroring and rotating the five main possibilities, a combination of all possible cases can be produced and registered in the look-up table. Tetrahedral splitting was carried out by proposing a generic 1:17 tetrahedral split, where the current geometry of a surface cut is represented by replacing the reference edge and face midpoints by the current intersection points, which are computed by the collision detection algorithm. By referencing the edge mid-nodes twice and the face mid-nodes three times, a pre-split tetrahedron composed of five parts can be produced. Although, the generic subdivision method produces acceptable results, it has some limitations. This include; producing hanging nodes that have no connection to neighbouring tetrahedral, which may lead to cracks in the model. This problem can be solved by splitting adjacent tetrahedral accordingly, but this has a trade-off between accuracy and computational load. Also, because splitting the tetrahedral is carried out after completing the cut, this may produce some visual inaccuracies and discontinuities.

Progressive Minimal cutting is another effective procedure proposed by Mor and Kanade (2000), which generates a minimal set of new elements that follows the trajectory of the scalpel. Developers of this system argue that the cutting algorithms introduced so far do not split the intersected element until the cut has been completed, which causes lag into the cutting process. They have implemented a method based on progressive cutting, where the subdivision is based on the geometry of the original element. In a progressive cutting procedure, a temporary subdivision of each partially cut element is encountered. Any temporary face interactions are updated for each partially cut tetrahedron. Then, the modified topology of the partially cut tetrahedron is checked for any changes. If the topology has changed, then a new set of temporary tetrahedral overwrites the old set. If the topology has not changed, then the temporary tetrahedral are updated by applying the new positions of the new face intersections. As soon as the cutting edge leaves a tetrahedron and the cut is finished, the temporary elements are then deleted and a final subdivision procedure is applied again. Eleven different combinations of intersected edges, faces, and temporary face intersections were generated to model the possible cut combinations. Mor and Kanade (2000) suggest that cutting can be affected by three factors,

- Cutting should occur along a free form path traced by the cutting tool.
- There should be no time lag between the movement of the scalpel and the resulting cut.
- The generated number of elements by cuts should be as small as possible, as this will affect the update rate significantly.

2.3.3 Relaxation. Relaxation defines the procedural computation of the physical behaviour of tissue after cutting.

2.3.4 Examples of cutting implementations. Outlined below are some implementations that have used the aforementioned cutting approaches and other application specific techniques.

A progressive cutting routine was used by Mor and Kanade (2000) to cut through tetrahedral models of soft tissue. Developers of this system claim that such technique increased accuracy and applicability of their simulator. Zhang et al (2002) also employed progressive cutting to subdivide the surface and produce interior structures that follow the path of the cutting device. Bielser et al (2003) also used progressive cutting approach to subdivide tetrahedral meshes. They have used a state machine to track the topology of tetrahedral and then control their progressive subdivision.

A combination of finite element deformation with interactive cutting was implemented by Nienhuys and Stappen (2000, 2001). An iterative algorithm was used to simulate linear deformation and cuts of tetrahedral meshes at real time with no pre-computation. The separation between the topological and geometric aspects is maintained, which made cutting interactive and fast.

Virtual freeform incisions on Finite Element tetrahedral meshes can be produced as suggested by Mazura and Seifert (1997). In their implementation, the user specifies a series of 3D points and a corresponding depth on the surface. Each pair of the succeeding points then defines the incision as a freeform shape.

A multi-resolution based cutting approach was introduced to ensure the required speed and to support dynamic changes of the topology at real time (Ganovelli et al, 2000, 2001). The idea is that changes in structure are only applied to parts of the model, such as cuts or lacerations. This system is based on tetrahedral decomposition of the space.

Bro-Nielsen et al (1998) have used mass-spring models to achieve reasonable cut results. They have accomplished this by applying a set of operations to individual triangles. A linear mass-spring model was used to model arteries and other tabular models as the backbone of a structure, which based on connected contours.

Cutting, tearing, and suturing can be simulated using the chainmail (Frisken-Gibson, 1999; Gibson et al, 1998). Cuts can be produced by breaking the connections between elements along the path of the virtual knife instrument. Intersection between the knife and the object are detected by scanning the knife space through the occupancy map and checking for collisions. Tearing happens when the distance between two elements is stretched beyond the allowable limit. When this limit is exceeded and cannot be compensated by the moving neighbours, then the connections are broken to simulate tearing. During suturing, elements along the path of the suturing instrument, which have missing links, are paired to their neighbours. These elements can be detected using the occupancy map within the vicinity of the suturing instrument.

Cotin et al (2000) have implemented real time deformations and cuts using linear elasticity. They have used three models do their simulations. The first model was built to make deformations, but no topological changes. The second was based on “tensor-mass” model, which allowed the simulation of deformations and cuts on small size meshes. A hybrid model based on the other two allowed the simulation of deformations and cuts on complex meshes.

Pflesser et al (1998, 2002) have developed a way of specifying, and modelling arbitrary shaped cut surfaces in volumetric models. These freeform cuts were represented within a voxel model together with partially volumetric effect. Using a grey-level gradient method allowed the calculation of accurate surface normals, which can be used to find sub-voxel localization of cut surfaces.

2.4 Task Specific Approaches

In this section we outline some systems, which have been built using approaches slightly different from the aforementioned techniques.

2.4.1 Karlsruhe endoscopic surgery simulator. Karlsruhe has developed an advanced endoscopic surgery simulator (Karlsruhe, 1997; Kuehnapfel et al, 1997) capable of producing highly realistic tissue behavior. In order to model such realistic behaviours, developers have adopted an approach, which is based on three different modelling elements. These are; the physical modelling, the geometric modelling, and the model interaction (Karlsruhe, 2001). To model the physical behaviour, a simple-element-system called “nodal net model” was employed. This model was implemented based on physical equations applied on discrete components. The result system with mass knots produces a system with differential equations, which can be solved in real time numerically. A B-Spline surface software was used in tissue (physical) modelling. Two methods were employed for pure data representation, a freeform surfaces method, and a direct output of polygonal nets. To connect the physical model with the geometric model, the mass knots were connected to the related control vertices of the polygonal net. Interaction involved three elements, collision detection, interaction management, and model modification. The deformable objects behaviour has to be dependant on the manipulation model hence instrument manipulations simulate actions such as, cutting, and grasping.

2.4.2 Endoscopic sinus surgery simulator. This endoscopic sinus surgery system was developed using the (NLM) Visible Male dataset (NLM, 2001) to model body tissue. The used dataset included MRI, CT, and colour cryoslice photographs. The rendered model has been developed through three stages, Segmentation, Surface Extraction, and Surface Simplification (Lockheedmartin, 2003).

Due to the high spatial resolution of the NLM cryoslice images and the difficulty in extracting details of such images, segmentation masks with unique colour for each image segment had to be produced. This permitted the developers to enhance the data by adjusting the segmentation masks. For surface extraction, developers had employed the marching cube algorithm from General Electric. For surface simplification process, an algorithm was used to reduce the high frequency in the polygonal mesh. In order to compute the surface structure in real-time, the polygons were classified into four groups, low, medium, high, and very high resolution. It has been noted that low and medium resolution polygons deliver acceptable results in real-time, so they were used for the proposed simulator. Textures were applied to the low fidelity data, and various anatomical structures were contorted, which have generated different pathological variants.

2.4.3 Gynaecology endoscopic simulator. Here we refer to the endoscopic simulator mentioned in §2.2.1. The modelling process of such system was divided into two stages, anatomical modelling and organ appearance modelling (Székely et al, 1999).

The Visible Human Female dataset from the National Library of Medicine (NLM, 2001) was selected to provide the data for constructing the simulator anatomical model. This choice has been made, because of the relatively high resolution and due to the consistency of the images provided. In order to produce 3D images from this standard 2D datasets, a segmentation system was used. Texturing was selected to model organ appearance, because it enhances realism and gives a cue to space perception. Also, texturing is crucial in modelling pathological tissue, which helps improve diagnostic skills. Organ specific base textures were produced using a texture analysis/synthesis procedure. A small texture sample is taken from a real image can be computed in any analysis stage. A 3D texture block is then modified until its second order statistics is close to the sampled texture. Organs then can be textured on the solid texture block.

3. COMPARISION

Object modelling approaches each have distinct advantages and disadvantages. Pure volumetric representations are ideal for true volume modelling, because they preserve the original data from the scanned images, but they require huge storage space and powerful software and hardware to render. Surface representations are easy to render and do not need powerful computers, but are an artistic impression of the body rather than a faithful representation of real data.

We have seen that the deformation approaches mentioned complement each other in many ways. For example, FEM can produce realistic and accurate models, but they are slow when manipulated. On the other hand, Mass Spring modes are relatively fast, but they lack accuracy and do not generate realistic images. Few simulations were built using the Chainmail method, but it is a promising candidate for future volumetric simulations. 3D linear elasticity is only valid for small deformations, which can be used mainly to simulate deformations associated with cuts. Linear elasticity can not be used for the modelling of all soft tissue materials, because most of these materials have a non-linear behaviour.

It can be seen that the intended aim of the simulating system may suggest which deformation approach to use. If the purpose of the simulation system focuses only on the visual part and ignores the time required to render that model, then FEM will be the ideal candidate. But if our focus is to carry out deformations in real time, then Mass-Spring method will be adopted. If we wanted volumetric simulations with greater speed computations, then chainmail method with volumetric representations are recommended.

With regard to cutting, we have seen how effective progressive cutting technique is in simulating cuts on tetrahedral surfaces, which suggests that this approach is a promising one. But this technique was only applied to tetrahedral surfaces with no deformations involved. A way of applying this method with deformable models will be an advantage. Also, applying progressive cutting on volumetric models, such as those used by Pflesser et al (1998, 2002) will be beneficial. No single approach was claimed to be the ideal solution. A combination of such techniques, especially when deformation and cutting involved, may provide a common ground for a universal solution, but this will require some time and effort to apply and test. We have also learned that there is a trade-off between accuracy and speed in simulating cuts, using the aforementioned implementations. Individual systems tackled portions of the problem, not the whole. It is clear that simulating cut on tissue is still a research topic. Further investigations are required before we see a system that delivers such task with the required accuracy and at acceptable computational rates.

4. CURRENT LIMITATIONS AND FUTURE EXPECTATIONS

In the following section we will outline some of the current limitations of modelling techniques and in the next one we will highlight some of their future expectations.

5.1 Current Limitations

Despite all the satisfactory results obtained by using the aforementioned modelling approaches, there still many limitations, which keeps medical simulation a research subject. Some of these limitations include,

- Modelling techniques used and to what degree they can be employed are governed by the required degree of complexity and realism and their relation to computational speed, storage, and access.
- Few open surgery simulators exist. This is because this type of surgery involves many tissue types with different nature and characteristics that can not be easily modelled using today's technology. Also, because this type of surgery is based on direct involvement of surgeons' hands which are versatile and hard to mimic.
- Various modelling approaches lend themselves to particular application, because of their characteristics, particularly in terms of appearance and rendering speed.

In summery we can say that volumetric modelling approaches based on real data sets are good when it comes to realistically modelling the human organs, but there is a big trade-off between what looks real and how long it takes to display and manipulate. Also, there is no ideal modelling approach, for medical simulators. The type of simulator and its intended purpose and the required degree of fidelity play a major role in deciding which approach to use.

5.2 Future Expectations

The wide acceptance of simulators as tools for medical planning and training will require the defeat of some challenges, particularly in terms of model representation, deformation, and cutting, some of these challenges include,

- Develop further expertise and mechanism to address the problem of complexity and realism and their relation to computational speed and storage.
- Improvements in fidelity and accuracy of deformation, cutting, appearance, and haptic cues, particularly with respect to complex open surgery.

5. REFERENCES

D Bielser, P Glardon, M Teschner and M H Gross (2003), A State Machine for Real-Time Cutting of Tetrahedral Meshes, *Proceedings of Pacific Graphics (PG'03)*, October 8-10, Canmore, Alberta, Canada, pp. 377-386.

D Bielser and M H Gross (2000), Interactive simulation of surgical cuts, *In Proc. Pacific Graphics, IEEE Computer Society Press*, pp. 116-125.

D Bielser, V A Maiwald and M. H. Gross (1999), Interactive Cuts Through 3-Dimensional Soft Tissue, *Proc. of the Eurographics '99, Computer Graphics Forum*, **18**, 3, pp. C31–C38.

M Bro-Nielsen, D Helfrick, B Glass, X Zeng and H Connacher (1998), VR simulation of abdominal trauma surgery, *In Proc. MMVR'98, IOS Press*, San Diego, pp. 117-123.

Colorado University (2004), Real-Time Visually and Haptically Accurate Surgical Simulation, Centre for Human Simulation, Electronic version: http://www.uchsc.edu/sm/chs/research/mmvr.html.

S Cotin, H Delingette and N Ayache (2000), A Hybrid Elastic Model allowing Real-Time Cutting, Deformations and Force-Feedback for Surgery Training and Simulation. *The Visual Computer,* **16**, 8, pp. 437-452.

S Cotin, H Delingette, and N Ayache (1999), Real-time elastic deformations of soft tissues for surgery simulation, *IEEE Transactions On Visualization and Computer Graphics*, **5**, 1, pp.62-73.

H Delingette (1998), Towards Realistic Soft Tissue Modelling in Medical Simulation, *Proc. IEEE: Special Issue on Surgery Simulation,* pp 512-523.

F Ganovelli, P Cignoni, C Montani and R Scopigno (2000), A Multiresolution Model for Soft Objects supporting interactive cuts and lacerations, *Computer Graphics Forum*, **19**, 3.

F Ganovelli, P Cignoni, C.Montani and R.Scopigno (2001), Enabling Cuts on Multiresolution Representation**,** The Visual Computer*, Springer International*, **17**, 5, pp.274-286.

Frisken-Gibson (1999), Using Linked Volumes to Model Object Collisions, Deformation, Cutting, Carving, and Joining*, IEEE Transactions on Visualization and Computer Graphics*, **5**, 4.

S Gibson, J Samosky, A Mor, C Fyock, E Grimson, T Kanade, R Kikinis, H Lauer, N McKenzie, S Nakajima, H Ohkami, R Osborne and A Sawada (1998), Simulating arthroscopic knee surgery using volumetric object representations, real-time volume rendering and haptic feedback, *Medical Image Analysis*, **2**, 2, pp.121-132.

S Gibson, C Fyock, E Grimson, T Kanade, R Kikinis, H Lauer, N McKenzie, A Mor, S Nakajima, H Ohkami, R Osborne, J Samosky and A Sawada (1998), Volumetric object modelling for surgical simulation, Medical Image Analysis, **2**, 2, pp. 121-132.

S Gibson and B Mirtich (1997), A survey of deformable modelling in computer graphics, *TR-97-19, Mitsubishi Electric Research Laboratory*, Cambridge, MA, USA.

S Gibson (1997), 3d chainmail: A fast algorithm for deforming volumetric objects, *In Proc. Symp. on Interactive 3D Graphics*, Providence, RJ, pp. 149-154.

Forschungszentrum Karlsruhe (2001): http://www-kismet.iai.fzk.de/TRAINER/mod_struc1.html.

Forschungszentrum Karlsruhe (1997): http://www-kismet.iai.fzk.de/TRAINER/mic_trainer1.html.

U Kuehnapfel, C Kuhn, M Huebner, H Krumm, H Maafl, and B Neisius, (1997), The karlsruhe endoscopic surgery trainer as an example for virtual reality in medical education, *In Minimally Invasive Therapy and Allied Technologies*, Blackwell Science Ltd, **6**, pp.122-125.

Mitsubishi (2001), Electronic Research Laboratories, Knee Arthroscopy Simulation Using Volumetric Knee Models, Electronic version: http://www.merl.com/projects/kneesystem2/.

L D Landau and E M Lifschitz (1986), Theory of Elasticity, Butterworth.

Lockheedmartin (2003), Medical Simulation, ENT Surgical Simulator, Mid-Term Report, Electronic version: http://www.lockheedmartin.com/akron/busdev/sim&trng/medsim/report/midterm.htm.

A Mazura and S Seifert (1997), Virtual cutting in medical data, *In [MWHS97]*, pp.420-429.

A Mor and T Kanade (2000), Modifying Soft Tissue Models: Progressive Cutting with Minimal New Element Creation, *Proceedings of Medical Image Computing and Computer-Assisted Intervention - MICCAI* ,**1935**, pp. 598-607.

The National Library of Medicine (NLM) (2003), The Visible Human Project, Electronic version: http://www.nlm.nih.gov/research/visible/visible_human.html.

The National Library of Medicine (NLM) (2001), The Visible Human Project, Electronic version: http://www.nlm.nih.gov/research/visible/getting_data.html.

H W Nienhuys and A F van der Stappen (2001), Supporting cuts and finite element deformation in interactive surgery simulation, *Proceedings Medical Image Computing and Computer Assisted Intervention*.

H W Nienhuys and A F van der Stappen (2000), Combining finite element deformation with cutting for surgery simulations, In d.A. Sousa & J.C. Torres (Eds.), *EuroGraphics 2000 Short Presentation,* Interlaken, Zwitserland, pp. 143-152.

B Pflesser, A Petersik, UTiede, K Höhne and R Leuwer (2002), Volume cutting for virtual petrous bone surgery, *Computer. Aided Surgery,* **7**, 2, pp.74-83.

B Pflesser, U Tiede, and K H Höhne (1998), Specification, Modelling and Visualization of Arbitrarily Shaped Cut Surfaces in the Volume Model, *Medical Image Computing and Computer-Assisted Intervention*, pp.853-860.

A Radetzky and A Nürnberger (2002), Visualization and Simulation Techniques for Surgical Simulators Using Actual Patient's Data, *Artificial Intelligence in Medicine,* **26**,3, pp. 255-279.

A Radezky, M Rudolph, S Starkie, B Davies and M Auer (2000), ROBO-SIM: A Simulator for Minimally Invasive Neurosurgery using an Active manipulator, *Hasman et al. Amsterdam, Proc. of MIE2000*, *IOS Press, Studies in Health Technology and Informatics*, pp.1165-1169.

K D Reinig, C G Rush, H L Pelster, V M Spitzer and J A Heath (1996), Real-Time Visually and Haptically Accurate Surgical Simulation, *Proceedings of Medicine Meets Virtual Reality: Health Care in the Information Age,* IOS Press, pp 542-546.

G Székely, M Bajka, C Brechbühler, J Dual, R Enzler, U Haller, J Hug, R Hutter, N Ironmonger, M Kauer, V Meier, P Niederer, A Rhomberg, P Schmid, G Schweitzer, M Thaler, V Vuskovic, and G Tröster, (1999), Virtual Reality Based Surgery Simulation for Endoscopic Gynaecology, *Medicine Meets Virtual Reality (Proceedings of MMVR'99)*, Studies in Health Technology and Informatics 62, IOS Press, Amsterdam, **62**, pp. 351-357.

C Wagner, M Schill, and R Manner (2002), Collision Detection and Tissue Modelling in a VR-Simulator for Eye Surgery, *ACM International Conference Proceeding Series***,** *Proceedings of the workshop on Virtual environments*, pp.27-36.

H Zhang, S Payandeh, and J Dill (2002), Simulation of Progressive Cutting on Surface Mesh model, *DRAFT6-08 Sept02*.

O C Zienkiewicz and R L Taylor (2000), The Finite Element Method Volume 1: The Basics, Butterworth Heinemann.

Memory assessment using graphics-based and panoramic video virtual environments

A A Rizzo[1,2], L Pryor[3], R Matheis[4], M Schultheis[4], K Ghahremani[1] and A Sey[3]

[1]Institute for Creative Technologies, University of Southern California,
13274 Fiji Way, Suite 600, Marina del Rey, CA. 90292, USA

[2]School of Gerontology, University of Southern California
3715 McClintock Ave. MC-0191, Los Angeles, CA. 90089-0191, USA

[3]Annenberg School for Journalism, University of Southern California, Los Angeles, Calif. 90089-0281
[4]Kessler Medical Rehabilitation Research & Education Corp. (KMRREC)
1199 Pleasant Valley Way, West Orange, NJ 07052, USA

arizzo@usc.edu, lpryor@usc.edu, mschultheis@kmrrec.org

ABSTRACT

Virtual Reality (VR) technology offers new options for neuropsychological assessment and cognitive rehabilitation. If empirical studies demonstrate effectiveness, virtual environments (VEs) could be of considerable benefit to persons with cognitive and functional impairments due to traumatic brain injury, neurological disorders, learning disabilities and other forms of Central Nervous System (CNS) dysfunction. Testing and training scenarios that would be difficult, if not impossible, to deliver using conventional neuropsychological methods are now being developed that take advantage of the assets available with VR technology. These assets include the precise presentation and control of dynamic multi-sensory 3D stimulus environments, as well as advanced methods for recording behavioral responses. When combining these assets within the context of functionally relevant, ecologically valid VEs, a fundamental advancement emerges in how human cognition and functional behaviour can be assessed and rehabilitated. This paper focuses on the results of two studies that investigated memory performance in two VEs having varying levels of functional realism. Within these VEs, memory tests were designed to assess performance in a manner similar to the challenges that people experience in everyday functional environments. One VE used a graphics based simulation of an office to test object memory in persons with TBI and healthy controls and found that many TBI subjects performed as well as the control group. The other study compared healthy young persons on their memory for a news story delivered across three different display formats, two of which used a 360-Degree Panoramic Video environment. The results of this "in progress" study are discussed in the context of using highly realistic VEs for future functional memory assessment applications with persons having CNS dysfunction.

1. INTRODUCTION

The assessment and rehabilitation of cognition has received considerable attention in the neuropsychological (NP) literature and such research has demonstrated cognitive impairment to be quite common in Traumatic Brain Injury (TBI) (Levin, Gary, Eisenberg, 1990) and other forms of central nervous system (CNS) dysfunction. Such impairments in cognitive functioning include attention (Litvan, Grafman, Vendrell & Martinez 1998), information processing abilities (Diamond, DeLuca, Kim & Kelley, 1997), and memory functioning (Rosenthal & Ricker, 2000; Brassington & Marsh, 1998). Memory is one of the most consistently impaired functions identified in these populations, with current prevalence rates ranging from 54% to 84% in a TBI population (McKinlay & Watkiss, 1999). In addition, studies have indicated that deficits in memory functioning are a major factor in one's ability to maintain meaningful employment following TBI (McKinlay & Watkiss, 1999). Given the relationship between memory abilities and employment status/quality of life, the assessment of functionally relevant memory performance is of vital importance for persons with CNS dysfunction. Advances in functional memory assessment could serve to identify relevant areas of preserved memory strengths, as well as impairments, that could support the creation of interventions that aim to facilitate a return to gainful employment and enhance quality of life.

Traditionally, the mainstay of learning and memory assessment has been NP testing. However, traditional cognitive assessment (and rehabilitation) methods have been criticized as limited in the area of *ecological* validity, that is, the degree of relevance or similarity that a test or training system has relative to the "real" world, and in its value for predicting or improving "everyday" functioning (Neisser, 1978). Adherents of this view challenge the usefulness of traditional psychometric tests for measuring and addressing the complex integrated functioning that is required for successful performance in the real world. A primary strength that Virtual Reality (VR) offers cognitive assessment and rehabilitation is in the creation of simulated functional environments in which performance can be tested and trained in a systematic fashion. By designing virtual environments that not only "look like" the real world, but actually incorporate challenges that require functional behaviours similar to the real world, the ecological validity of assessment and rehabilitation methods could be enhanced. As well, within a VE, the experimental control required for rigorous scientific analysis and replication can still be maintained within simulated contexts that embody the complex challenges found in naturalistic settings. Thus, VR derived results could have greater predictive validity and clinical relevance for quantifying the challenges that patients face in the real world. The unique match between VR technology assets and the needs of various neuropsychological application areas has been recognized by a number of authors (Pugnetti et al., 1995; Rose, 1996; Rizzo, Schultheis, Kerns and Mateer, 2004) and an encouraging body of research has emerged (Rizzo, Buckwalter and van der Zaag, 2002).

This paper will present findings from two HMD VR memory assessment scenarios that possess varying levels of "realism" and place performance demands on participants that are more similar to "real-world" challenges compared to traditional list learning memory tests. The first study used a graphics-based simulation of a virtual office to compare object memory performance in subjects with TBI and healthy controls conducted at the KMRREC via a collaborative agreement with USC. The second study used a 360-degree Panoramic Video (PV) camera system to capture a news reporter presenting a two-minute news story from the streets of Los Angeles. With this system, users can observe pictorially accurate 360-degree video scenes of "real world" environments delivered via a head mounted display (HMD). Healthy young subjects were tested in terms of how much they could remember from the reported story (in similar fashion to the Wechsler Memory Scale III - Logical Memory Subtest), under varying levels of immersion. These results are serving as an initial feasibility test of this media format as a precursor to our future functional memory research using this system with persons having CNS dysfunction.

2. THE VIRTUAL OFFICE

2.1 Rationale for Development of the Virtual Office Application

Following on our previous work developing a Virtual Classroom for the assessment of attention processes in children with ADHD (Rizzo et al., 2004), we have created other scenarios (i.e., work situations, home environments, etc.,) using the same logic and approach to address cognitive/functional processes that are relevant for a range of other clinical populations. In this regard, we have now constructed a Virtual "Office" environment (see Figure 1) that evolved from expanding some of the basic design elements of the USC Virtual Classroom. This scenario was generally conceptualized as an "open platform" that could be used to study, test and train a variety of cognitive processes depending on the research question. Within this version of the virtual office, it is possible to place objects in strategic locations in the VE, remove them with a keystroke, and within this format, collect performance data on memory for objects in an environment that resembles a functional setting in everyday life.

2.2 Methods

2.2.1 Participants. The present study recruited 40 participants, 20 individuals with traumatic brain injury (TBI) and 20 healthy controls (HC), matched on age, sex, and education (See Table 1). All participants were between the ages of 18 and 55, were medically stable and had no significant psychiatric, neurological (i.e., other than TBI), or substance abuse history. TBI participants all sustained injuries classified as severe, with Glasgow Coma Scores (GCS) of 8 and/or loss of consciousness (LOC) for no less than 6 hours immediately following injury. Participants with significant visual disturbance (i.e., impeding ability to participate in the VE), aphasia, anomia, and or history of learning disorder were excluded from the study.

2.2.2 VE scenario and Testing Procedure. Three categories of measures were administered. First, general questionnaires were used to solicit past medical and psychiatric histories, as well as, demographic information. Second, a traditional NP battery was administered to assess all major domains of neurocognitive

©2004 ICDVRAT/University of Reading, UK; ISBN 07 049 11 44 2

functioning. Finally, all participants were administered the "Virtual Office" memory task. The current "Virtual Office" task is a computer-generated environment designed to simulate a generic office setting. The system equipment included a Dell Inspiron 8100 laptop, a 5th Dimension Technologies (5DT) 800 Series Head Mount Display (HMD), and a Flock of Birds position and orientation tracking system. Participants "entered" the virtual office by placing the HMD display on their head. The HMD could be flipped up and down. When HMD is in the up position, the participant was removed from the VE and could see everything in the "real world", with no view of the virtual office. In the down position, the participants' view of the "real world" was occluded and his/her visual field was filled with images of the virtual office environment.

Figure 1. *Scenes from the original Virtual Office and a new version currently under construction.*

Upon entering the virtual office, participants appeared to be seated at a desk and had a head-level view of the entire office space. The office task was programmed to include sixteen target items to be remembered. The sixteen target items consisted of both eight common office items (e.g., notepad) and eight uncommon items (e.g., stop sign). This selection of objects precluded participants from inflating their scores simply by recalling typical office items and allowed us to assess differences in memory based on object novelty effects. The VR task itself was a test of learning and memory. The task began when participants entered the virtual office by wearing the HMD. While "inside" the office, the participants received an audio-guided tour that named the target items to be remembered, each time in a different order. Participants then exited the VE (i.e., by removing HMD) and were asked to recall from memory all target items seen in the VR office. This procedure of entering and exiting the office constituted one learning trial. All participants continuously received learning trials until they could recall all target items across two consecutive trials (i.e., the learning criterion) or until the maximum of twelve trials was reached. After a 30-minute delay and again after a 24-hour delay, participants again were asked to recall as many target items as possible. Following informed consent procedures, participants were administered both the NP battery and the VR task. Order of completion of these tasks was counterbalanced to control for fatigue. All participants were contacted by phone approximately twenty-four hours following test administration to collect long-term recall data. Specifically, they were asked to recall all the items seen in the virtual office on the previous day.

2.2 Results

Among the 20 participants with TBI, 20% (n = 4) were unable to meet the VR Office task learning criterion, whereas 100% of the HC participants met the criterion. Thus, three groups were identified: individuals with TBI who met the criterion (TBI-MET, n = 16), individuals with TBI who did not meet the criterion (TBI-NOT MET, n = 4), and HC participants (n = 20). (See Table 1).

Initial target acquisition. To examine group differences in initial acquisition of target stimuli, mean number of trials to criterion were compared between the TBI-MET and HC groups. TBI-NOT MET participants were excluded as, by definition, they did not meet the learning criterion. Results indicated that the TBI-MET and HC groups were nearly identical in the number of trials required to learn the 16-items presented, $F(1, 34) = 0.00$, $p = 0.98$). Specifically, the TBI-MET group required an average of 3.94 trials (SD = 1.61; range,2-8), whereas the HC group required an average of 3.95 trials (SD = 1.91; range, 2-10) to meet criterion.

Target recall. Virtual office recall performance among the three groups was compared with a 3(Group) X 2(Delay) repeated measures analysis of variance (ANOVA). There was a significant main effect for Group ($F(1, 29) = 43.2$, $p < 0.001$), and post hoc Tukey tests revealed that the TBI-NOT MET group recalled significantly fewer items (M = 8.25, SD = 4.57) than both the TBI-MET group (M = 14.31, SD = 1.70) and the HC group (M = 15.35, SD = 1.04). The TBI-MET and HC groups did not differ significantly in overall item recall. There also was a significant main effect for Delay ($F(1, 29) = 43.2$, $p < 0.001$), where as expected when data was collapsed across the 3 groups, significantly fewer items were recalled at the 24-hour delay (M = 7.85, SD = 0.72) than at the 30-minute delay (M = 12.67, SD = 0.41). Notably, the interaction of Delay and Group also was statistically significant ($F(2, 29) = 4.23$, $p < 0.05$), where performance of the TBI-NOT MET group varied over time. Interestingly, while the TBI-NOT MET group recalled significantly fewer items than the other 2 groups following a 30-minute delay, this difference disappeared following 24-hour delay, as the groups did not differ significantly.

Table 1. *Demographic Characteristics Organized By Group.*

	TBI-MET (*n* = 16)	TBI-NOT MET (*n* = 4)	HC (*n* = 20)
Gender			
Women	7 (44%)	2 (50%)	11 (55%)
Men	9 (56%)	2 (50%)	9 (45%)
Age (in years)			
Mean (*SD*)	37.6 (9.1)	39.5 (11.8)	31.6 (12.4)
Education (in years)			
Mean (*SD*)	13.3 (2.1)	14.0 (2.8)	15.2 (2.4)
Marital Status			
Single	13 (81%)	3 (75%)	13 (65%)
Married	3 (19%)	1 (25%)	7 (35%)
Ethnicity			
Caucasian	15 (94%)	3 (75%)	18 (90%)
African American	0	1 (25%)	1 (5%)
Hispanic	1 (6%)	0	1 (5%)
Employment Status			
Employed	10 (63%)	2 (50%)	13 (65%)
Unemployed	6 (37%)	2 (50%)	7 (35%)
Time LOC* - days	12.7 (15.3)	29.5 (20.5)	--
Time Post-Injury (yr.)	7.5 (5.4)	6.5 (4.2)	--

Note. *LOC = Loss of consciousness.

Traditional Neuropsychological Test Summary. Results of neuropsychological testing revealed significant differences in performance between the HC and TBI groups. Specifically, when compared to the HC group both TBI groups demonstrated significantly lower psychometric intelligence, psychomotor speed, auditory short-term attention, numeric sequencing, executive functioning, visual scanning, confrontation naming, and verbal fluency ($p < 0.05$). Learning and memory was assessed with the California Verbal Learning Test (CVLT) (Delis et al., 1987). Significant group differences also were observed on this verbal list-learning task, as the HC group learned a significantly greater number of words after 5 learning trials and recalled a significantly greater number of words after a 30-minute delay ($p < 0.001$). Overall, TBI groups demonstrated significantly poorer performance than HCs using these traditional measures.

2.4 Conclusions

The present study compared a group of persons with TBI to a group of matched, healthy controls using standard NP tests and on a VR-based memory assessment instrument. Interestingly, a large percent of participants with brain injury were statistically equivalent to healthy controls in their ability to acquire target items during the learning trials of the virtual office task. Furthermore, recall of the target items at both 30 minutes and at 24 hours was not significantly different between some of these participants and the HC group. These findings indicate that many TBI participants failed to demonstrate impaired acquisition and retrieval when using the VR Office as a measure of object memory.

This finding is at variance with the earlier work of DeLuca, Schultheis, Madigan, Christodoulou, and Averill (2000), who conducted an analogous study and observed TBI participants to have significantly worse initial acquisition and retrieval of target items than a group of matched, healthy controls using a commonly employed verbal list learning task, The Selective Reminding Task (SRT) (Buschke, 1973). One important way in which these two tasks differ is that the VR Office provides a rich context in which visual target stimuli are presented. This context is not available in the SRT or in other traditional verbal list-learning tasks and therefore a different memory process is being assessed. Similar results were found in the current study using the CVLT whereby participants appeared able to benefit from the presence of contextual visuospatial cues, which might have enhanced initial encoding of target items. Specifically, participants were afforded the opportunity to visually associate target items with elements of the environment, as well as with other target items in the virtual office. In addition, during testing with VR numerous elements of memory (e.g., verbal, visual, spatial) are assessed in combination, thereby potentially improving scores beyond traditional measures that are typically used to assess components of memory in isolation. Indeed, this may have improved the quality with which target items were encoded and reduced the quantity of learning trials required to meet the learning criterion. Perhaps the availability of such contextual cues during testing with VR more closely mimics the affordances in the real-world compared to traditional verbal measures of memory. In this regard, we could speculate that either the VR task has limited utility for isolating certain memory component impairments, or that the context provided in this type of VR assessment revealed preserved integrated functional memory ability that would be underestimated if memory assessment was limited to the types of verbal learning tests that are commonly employed in traditional NP assessment. The initial results from this research are highly suggestive of a future research direction that could have value for broadening our understanding of everyday memory and for perhaps guiding rehabilitative strategies with better input on preserved functioning.

3. PANORAMIC VIDEO VR MEMORY TEST

3.1 Rationale for Panoramic Video VR Memory Test

Recent advances in Panoramic Video (PV) camera systems have produced new methods for the creation of virtual environments (James, 2001). With these systems, users can capture, playback and observe pictorially accurate 360-degree video scenes of "real world" environments. When delivered via an immersive head mounted display (HMD), an experience of presence within these captured scenarios can be supported in human users. This is in sharp contrast to the constrained delivery and passive viewing of television and video images that have been the primary mode for providing humans with a "virtual eye" into distant times and locations over the last fifty years. Along with traditional computer graphics (CG) based virtual environments, PV overcomes the passive and structured limitations of how imagery is presented and perceived. The recent convergence of camera, processing and display technologies make it possible for a user to have control and choice in their viewing direction. As opposed to mouse and keyboard methods for interacting with flat screen panoramic content, users can more intuitively observe PV content via natural head movement within an HMD. Users of PV become virtual participants immersed in the observed scene, creating a new dimension in the way people perceive imagery within these types of VEs. However, when compared with CG-based VEs, PV has some limitations regarding functional interactivity. Whereas users operating within a CG-based VE scenario are usually capable of both 6DF navigation, and interaction with rendered objects, PV immersion allows mainly for observation of the scene from the fixed location of the camera with varying degrees of orientation control (i.e. pitch, roll and yaw). In spite of this limitation, the goals of certain application areas may be well matched to the assets available with this type of PV image capture and delivery system. One potential clinical application area in the use of PV content for creating standardized tests or training tools for addressing cognitive function within an ecologically enhanced real-world VE.

3.2 Brief system overview and technical description

Panoramic image acquisition is based on mosaic approaches developed in the context of still imagery. Mosaics are created from multiple overlapping sub-images pieced together to form a high resolution, panoramic, wide field-of-view image. Viewers often dynamically select subsets of the complete panorama for viewing. Several panoramic video systems use single camera images (Nayar, 1997), however, the resolution limits of a single image sensor reduce the quality of the imagery presented to a user. While still image mosaics and panoramas are common, we produce high-resolution panoramic video by employing an array of five video cameras viewing the scene over a combined 360-degrees of horizontal arc. The cameras are arrayed to look at a five-facet pyramid mirror. The images from neighbouring cameras overlap slightly to facilitate their merger. The camera controllers are each accessible through a serial port so that a host computer can save and restore camera settings as needed. The complete camera system (see Figure 2) is available from FullView, Inc. (FullView, 2001).

Figure 2. *FullView Panoramic Camera*

The five camera video streams feed into a digital recording and playback system that we designed and constructed for maintaining precise frame synchronization. All recording and playback is performed at full video (30Hz) frame rates. The five live or recorded video streams are digitized and processed in real time by the computer system. The camera lens distortions and colorimetric variations are corrected by our software application and a complete panoramic image is constructed in memory. With five cameras, this image has over 3000x480 pixels. From the complete image, one or more scaled sub-images are extracted for real-time display in one or more frame buffers and display channels. Figure 3 shows an example of the screen output with a full 360° still image extracted from the video.

The camera system was designed for viewing the images on a desktop monitor. With a software modification provided by FullView Inc. (FullView, 2001), we were able to create an immersive viewing interface using a head-mounted display (HMD). A single window with a resolution of 800x600 is output to the HMD worn by a user. A real-time inertial orientation tracker (Intersense, 2001) is fixed to the HMD to sense the user's head orientation. The orientation is reported to the viewing application through an IP socket, and the output display window is positioned (to mimic pan and tilt) within the full panoramic image in response to the user's head orientation. View control by head motion is a major contributor to the sense of immersion experienced by the user. It provides the natural viewing control we are accustomed to without any intervening devices or translations.

3.3 Panoramic Video Memory Test Scenario Design and Method

This project was conducted in collaboration with the USC Annenberg School of Journalism as part of a research effort aiming to investigate the use of Panoramic Video for viewing newsworthy content, as well as its impact on memory for the verbal content of a news story compared to traditional viewing methods. The news "stimuli" consisted of a female reporter presenting a story from a fixed position on a street in downtown Los Angeles. The news story involved a two-minute report on issues regarding the "homeless" in Los Angeles. The camera was positioned in the middle of a street in the midst of an array of tents and makeshift living quarters on the sidewalk. In addition to seeing the reporter, the 360-degree PV content contained the imagery and sounds of the surroundings with many homeless individuals going about their day-to-day activities in this area (See Figure 4). With this captured content we compared memory performance with a between groups design across three different viewing format conditions. The three conditions were:

Single Frame Condition (C1) subjects viewed the 2-minute news story in a "traditional" single-frame viewing format on a computer monitor. This group of participants had access only to the single frame field of view containing the news reporter's standing delivery of the story, as is commonly seen in a standard on-the-scene reporting approach presented on a television news broadcast.

Flat screen Panoramic Condition (C2) participants had access to view the complete 360-degree arc of the environment from where the C1 news story was reported. They had access to and viewed the 360-degree arc on a computer monitor using an inertial orientation tracker mounted on a disc to freely navigate around the panoramic arc. C2 subjects also heard the *exact same verbal delivery and had access to the same audiovisual presentation of the reporter* as presented in C1, since the C1 story was actually a flat panel extracted from the full panoramic 360-degree arc used in C2.

HMD Panoramic Condition (C3) participants viewed the exact same 360-degree arc of the news story environment that was available to the C2 group, but from within an orientation tracked HMD. This system updated the video image in the display in real time as the subject turned their head. This allowed the participant to view the scene as they would if they were at the site of the news story and to have free choice to observe the panoramic scene from any perspective within the 360-degree arc using head turning movements as they would under normal real world viewing conditions.

Figure 3. *360-degree PV image extracted from video footage taken at the Los Angeles Coliseum*

Figure 4. *Traditional Viewing Range (L) vs. Panoramic Viewing Range (R)*

Currently, 16 unimpaired research participants have been tested in each condition (avg. age=20y/o). Due to research participant acquisition challenges, the total sample reported on in this paper consists of 40 females and 8 males. The study is ongoing with a total sample of 96 anticipated, with equal gender representation expected. Participants were tested on free recall, cued recall and multiple choice recognition for the auditory content immediately following presentation of the news story, and in a delay condition (1 week later). This design allowed for comparison of memory across groups on immediate acquisition/recall/recognition of the story content and on long-term incidental recall/recognition retrieval. Initial analysis of the memory data with the existing sample is presented in this paper. Results to be presented at the conference will include head tracking analyses as a measure of exploratory scanning behaviour and its influence on memory, and results from the Presence Questionnaire (Witmer and Singer, 1998) to determine the relevance of this intervening variable as a factor in memory performance.

3.4 Panoramic Video Study Results and Conclusions

As seen in Table 2, single frame viewing produced significantly better memory than *both* PV conditions only for Immediate Recall. Single frame viewing was also shown to promote better immediate recognition memory than the flat screen panoramic condition. Memory performance was found to be equivalent for all conditions for immediate cued recall and for all delayed memory variables.

Table 2. *Memory Results.*

	Single Frame	Flat screen Panoramic	HMD Panoramic	Sig.
Recall Means (*SD*)				
Immediate	11.81 (4.7)	8.81 (4.6)	8.44 (2.8)	($F(2, 45) = 3.2$, $p < 0.05$)£
1-Week Delayed	9.94 (4.2)	7.75 (2.9)	7.44 (2.5)	NS
Cued Recall Means (*SD*)				
Immediate	5.44 (2.0)	4.38 (2.2)	3.94 (1.7)	NS
1-Week Delayed	6.19 (2.2)	4.69 (1.9)	5.0 (2.0)	NS
Recognition Means (*SD*)				
Immediate	8.25 (1.5)	6.56 (2.2)	7.31 (2.0)	($F(2, 45) = 3.1$ $p < 0.05$)*
1-Week Delayed	7.63 (1.8)	6.56 (1.9)	6.94 (1.4)	NS

£Post Hoc difference significant between Single Frame and both Panoramic Conditions combined.
*Post Hoc difference only significant between Single Frame and Flat screen Panoramic Condition.

One of our initial working hypotheses was that the sense of "being there" or "presence" would be enhanced in the HMD Panoramic Condition via the use of an immersive HMD and that this added engagement would increase *long term* recall by providing better contextual retrieval cues that leverage episodic memory processes. However, both Panoramic Conditions also provide the participant with additional information beyond the auditory content of the reported news story in the single frame condition. This occurs in the form of the real world activity that transpired in the full 360-Degree field of view that the user has available to explore. Whether this added cognitive load serves as a distraction that limits memory for the audio content or provides contextual cues that support memory retrieval is one of the open questions that this research will address via more refined analyses relating memory to presence and head movement. Data from those analyses from this preliminary sample will be available at the conference. One of the aims of this study was to determine if panoramic video VR would have added-value as a tool for testing everyday functional memory in clinical populations. These preliminary results suggest that in an unimpaired young sample, the added information available in the panoramic conditions did not impair long-term memory retrieval although initial recall was impacted. Future research with participants having CNS dysfunction will be conducted to determine if (unlike the unimpaired groups tested thus far) the added cognitive load that may exist in such Panoramic VR scenarios will serve to impair long-term memory performance. This could serve as the basis for a measurement tool that could provide a more relevant assessment of clinical impairments in everyday functional memory.

4. CONCLUSION

The projects briefly summarized in this paper reflect our view that VR technology offers assets that could potentially improve the reliability and validity of methods used in the areas of neuropsychological assessment and rehabilitation. The key elements for this exist in VR's capacity for consistent delivery of complex dynamic test and distraction stimuli within the context of functionally relevant simulated settings. In this manner, VR allows for systematic assessment and rehabilitation within simulated "real-world" functional testing and training environments with an aim towards enhancing ecological validity. Such a merger between traditional analog methods with more functional/contextual approaches, if successful, could remedy some of the limitations found individually with these approaches and result in more effective prediction of real world performance. The studies presented in this paper illustrate two approaches toward testing cognitive processes in environments that possess attributes similar to challenges commonly found in everyday memory. However, the value of this form of assessment in relation to traditional component testing will only be determined via rigorous experimentation with an eye towards enhanced explanation and prediction of memory impairments in a patients' day-to-day functional environment. It may be found that such testing will both reveal preserved functioning that could be leveraged in rehabilitation programming as well as exposing impairments in more integrated functioning that would not be uncovered with basic component tests.

5. REFERENCES

J C Brassington, N V & Marsh (1998). Neuropsychological aspects of multiple sclerosis. *Neuropsychology Review,* **8**, pp. 43-77.

H Buschke (1973). Selective reminding analysis of memory and learning. *Journal of Verbal Learning and Verbal Behaviour,* 12, pp. 543-550.

D C Delis, J H Kramer, E Kaplan & B A Ober (1987). *California Verbal Learning Test, Adult Research Edition.* New York: The Psychological Corporation.

J DeLuca, M T Schultheis, N Madigan, C Christodoulou & A Averill (2000). Acquisition versus retrieval deficits in traumatic brain injury: Implications for memory rehabilitation. *Archives of Physical Medicine and Rehabilitation*, **81**, pp. 1327-1333.

B J Diamond, J DeLuca, H Kim & S M Kelley (1997). The question of disproportionate impairments in visual and auditory information processing in multiple sclerosis. *Journal of Clinical & Experimental Neuropsychology.* **19**, pp. 34-42.

FullView.com Inc. (2003). Retrieved February 15, 2003, from www.fullview.com

Intersense Inc. (2003). Retrieved February 15, 2003, from www.isense.com

M S James (2001). 360-Degree Photography and Video Moving a Step Closer to Consumers. Retrieved March 23, 2001, from http://abcnews.go.com/sections/scitech/CuttingEdge/cuttingedge010323.html

H S Levin, H E Gary & H M Eisenberg (1990). Neurobehavioral outcome 1 year after severe head injury. *Journal of Neurosurgery.* **73**, pp. 669-709.

I Litvan, J Grafman, P Vendrell & J M Martinez (1988). Slowed information processing in multiple sclerosis. *Archives of Neurology.* **45**, pp. 281-285.

W McKinlay & A J Watkiss (1999). *Cognitive and Behavioral Effect of Brain Injury*. 3rd ed. F.A. Davis: Philadelphia, pp. 74-86.

S K Nayar (1997). Catadioptric Omnidirectional Camera, Proc. of IEEE Computer Vision and Pattern Recognition, (CVPR).

U Neisser (1978). Memory: what are the important questions? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical Aspects of Memory*. London: Academic Press. 3-24.

L Pugnetti, L Mendozzi, A Motta, A Cattaneo, E Barbieri and S Brancotti (1995). Evaluation and retraining of adults' cognitive impairments: Which role for virtual reality technology? *Computers in Bio. and Medicine,* **25**, *2*, pp.213-227.

A Rizzo, M Schultheis, K Kerns & C Mateer (2004). Analysis of Assets for Virtual Reality Applications in Neuropsychology. *Neuropsychological Rehabilitation.* **14**(1/2) 207-239.

A Rizzo, D Klimchuk, R Mitura, T Bowerly, C Shahabi & J G Buckwalter (2004). Diagnosing Attention Disorders in a Virtual Classroom. *IEEE Computer*. **37** (5), 87-89.

A Rizzo, J G Buckwalter, and C van der Zaag (2002), Virtual environment applications in clinical neuropsychology. In *The Handbook of Virtual Environments,* (K Stanney, Ed.), Erlbaum Publishing, New York. pp. 1027-1064.

F D Rose (1996). *Virtual reality in rehabilitation following traumatic brain injury,* In Proceedings of the European Conference on Disability, Virtual Reality and Associated Technology (P Sharkey, Ed.), Reading UK: The University of Reading, pp. 5-12.

M Rosenthal & J H Ricker (2000). Traumatic brain injury. In *Handbook of Rehabilitation Psychology.* (R Frank & T Eliot, Eds.), American Psychological Association: Washington, D.C.

D Wechsler (1997). Manual: Wechsler Memory Scale - III (WMS-III). San Antonio: Psychological Corporation.

B G Witmer and M J Singer (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*. 7, 225-240.

ICDVRAT 2004

Author Index

Author Index